GAME CHANGERS: ARTIFICIAL INTELLIGENCE PART I

HEARING

BEFORE THE

SUBCOMMITTEE ON
INFORMATION TECHNOLOGY

OF THE

COMMITTEE ON OVERSIGHT
AND GOVERNMENT REFORM
HOUSE OF REPRESENTATIVES

ONE HUNDRED FIFTEENTH CONGRESS

SECOND SESSION

FEBRUARY 14, 2018

Serial No. 115–65

Printed for the use of the Committee on Oversight and Government Reform

Available via the World Wide Web: http://www.fdsys.gov
http://oversight.house.gov

U.S. GOVERNMENT PUBLISHING OFFICE

30–296 PDF WASHINGTON : 2018

COMMITTEE ON OVERSIGHT AND GOVERNMENT REFORM

Trey Gowdy, South Carolina, *Chairman*

John J. Duncan, Jr., Tennessee
Darrell E. Issa, California
Jim Jordan, Ohio
Mark Sanford, South Carolina
Justin Amash, Michigan
Paul A. Gosar, Arizona
Scott DesJarlais, Tennessee
Blake Farenthold, Texas
Virginia Foxx, North Carolina
Thomas Massie, Kentucky
Mark Meadows, North Carolina
Ron DeSantis, Florida
Dennis A. Ross, Florida
Mark Walker, North Carolina
Rod Blum, Iowa
Jody B. Hice, Georgia
Steve Russell, Oklahoma
Glenn Grothman, Wisconsin
Will Hurd, Texas
Gary J. Palmer, Alabama
James Comer, Kentucky
Paul Mitchell, Michigan
Greg Gianforte, Montana

Elijah E. Cummings, Maryland, *Ranking Minority Member*
Carolyn B. Maloney, New York
Eleanor Holmes Norton, District of Columbia
Wm. Lacy Clay, Missouri
Stephen F. Lynch, Massachusetts
Jim Cooper, Tennessee
Gerald E. Connolly, Virginia
Robin L. Kelly, Illinois
Brenda L. Lawrence, Michigan
Bonnie Watson Coleman, New Jersey
Stacey E. Plaskett, Virgin Islands
Val Butler Demings, Florida
Raja Krishnamoorthi, Illinois
Jamie Raskin, Maryland
Peter Welch, Vermont
Matt Cartwright, Pennsylvania
Mark DeSaulnier, California
Jimmy Gomez,California

SHERIA CLARKE, *Staff Director*
WILLIAM MCKENNA, *General Counsel*
TROY STOCK, *Technology Subcommittee Staff Director*
SARAH MOXLEY, *Senior Professional Member*
SHARON CASEY, *Deputy Chief Clerk*
DAVID RAPALLO, *Minority Staff Director*

SUBCOMMITTEE ON INFORMATION TECHNOLOGY

Will Hurd, Texas, *Chairman*

Paul Mitchell, Michigan, *Vice Chair*
Darrell E. Issa, California
Justin Amash, Michigan
Blake Farenthold, Texas
Steve Russell, Oklahoma
Greg Gianforte, Montana

Robin L. Kelly, Illinois, *Ranking Minority Member*
Jamie Raskin, Maryland
Stephen F. Lynch, Massachusetts
Gerald E. Connolly, Virginia
Raja Krishnamoorthi, Illinois

CONTENTS

GAME CHANGERS: ARTIFICIAL INTELLIGENCE PART I

Wednesday, February 14, 2018

HOUSE OF REPRESENTATIVES,
SUBCOMMITTEE ON INFORMATION TECHNOLOGY,
COMMITTEE ON OVERSIGHT AND GOVERNMENT REFORM,
Washington, D.C.

The subcommittee met, pursuant to call, at 2:23 p.m., in Room 2154, Rayburn House Office Building, Hon. Will Hurd [chairman of the subcommittee] presiding.

Present: Representatives Hurd, Amash, Kelly, Lynch, Connolly, and Krishnamoorthi.

Also Present: Representative Massie.

Mr. HURD. The Subcommittee on Information Technology will come to order. And, without objection, the chair is authorized to declare a recess at any time.

Welcome to the first hearing in a series of hearings on artificial intelligence. This series is an opportunity for the subcommittee to take a deep dive into artificial intelligence. And today's hearing is an opportunity to increase Congress' understanding of artificial intelligence, including its development, uses, and the potential challenges and advantages of government adoption of artificial intelligence.

We have four experts on the matter whom I look forward to hearing from today. And in the next hearing we do, in March, I believe, we will hear from government agencies about how they are or should be adopting artificial intelligence into their operations, how they will use AI to spend taxpayer dollars wisely and make each individual's interactions with the government more efficient, effective, and secure.

It is important that we understand both the risks and rewards of artificial intelligence. And in the third hearing, in April, we will discuss the appropriate roles of both the public and private sectors as artificial intelligence matures.

Artificial intelligence is a technology that transcends borders. We have allies and adversaries, both nation-states and individual hackers, who are pursuing artificial intelligence with all they have, because dominance in artificial intelligence is a guaranteed leg up in the realm of geopolitics and economics.

At the end of this series, it is my goal to ensure that we have a clear idea of what it takes for the United States to remain the world leader when it comes to artificial intelligence. Thoughtful engagement by legislators is key to this goal, and I believe that this committee will be leaders on this topic.

So what is artificial intelligence? Hollywood's portrayal of artificial intelligence is not accurate. Instead, many of us are already using it every single day, from song recommendations in Spotify to digital assistants that tell us the weather.

And while these consumer applications are important, I am most excited about the possibility of using artificial intelligence in the government to defend our infrastructure and have better decision-making because of the analytics that artificial intelligence can run.

In an environment of tightening resources, artificial intelligence can help us do more for less money and help to provide better citizen-facing services.

I thank the witnesses for being here today and look forward to hearing and learning from you so that we can all benefit from the revolutionary opportunities AI provides us.

As always, I am honored to be exploring these issues in a bipartisan fashion, I think the IT Subcommittee is a leader on doing things in a bipartisan way, with my friend and ranking member, the Honorable Robin Kelly from the great State of Illinois.

Ms. KELLY. Thank you. Welcome to the witnesses. Thank you, Chairman Hurd, and welcome to all of our witnesses today, and Happy Valentine's Day.

Artificial intelligence, or AI, has the capacity to improve how society handles some of its most difficult challenges.

In medicine, the use of AI has the potential to save lives and detect illnesses early. One MIT study found that using machine-learning algorithms reduced human errors by 85 percent when analyzing the cells of lung cancer patients. And earlier this month, Wired magazine reported hospitals have now begun testing software that can check the images of a person's eye for signs of diabetic eye disease, a condition that if diagnosed too late can result in vision lost.

In some communities around the country, self-driving cars are already operating on the road and highways. That makes me nervous. Investment by major car companies in self-driving cars makes it increasingly likely that they will become the norm, not the exception on our Nation's roads.

But there is a lot of uncertainty revolving around artificial intelligence. AI is no longer the fantasy of science fiction and is increasingly used in everyday life. As the use of AI expands, it is critical that this powerful technology is implemented in an inclusive, accessible, and transparent manner.

In its most recent report on the future of AI, the National Science and Technology Council issued a dire assessment of the state of diversity within the AI industry. The NSTC found that there was a, quote, "lack of gender and racial diversity in the AI workforce," and that this, quote, "mirrors the lack of diversity in the technology industry and the field of computer science generally." According to the NSTC, in the field of AI improving diversity, and I quote, "is one of the most critical and high priority challenges."

The existing racial and gender gaps in the tech industry add to the challenges the AI field faces. Although women comprise approximately 18 percent of computer science graduates in the Nation, only 11 percent of all computer science engineers are female.

African Americans and Hispanics account for just 11 percent of all employees in the technology sector, despite making up 27 percent of the total population in this country.

Lack of AI workforce diversity can have real cost on individuals' lives. The increasing use of AI to make consequential decisions about people's lives is spreading at a fast rate. Currently, AI systems are being used to make decisions by banks about who should receive loans, by government about whether someone is eligible for public benefits, and by courts about whether a person should be set free.

However, research has found considerable flaws and biases can exist in the algorithms that support AI systems, calling into question the accuracy of such systems and its potential for unequal treatment of some Americans. For AI to be accurate, it requires accurate data and learning sets to draw conclusions. If the data provided is biased, the conclusions will likely be biased. A diverse workforce will likely account for this and use more diverse data and learning sets.

Within the industry, the use of black box algorithms are exacerbating the problems of bias. Two years ago, ProPublica investigated the use of computerized risk prediction tools that were used by some judges in criminal sentencing and bail hearings.

The investigation revealed that the algorithm the systems relied upon to predict recidivism was not only inaccurate, but biased against African Americans who were, quote, "twice as likely as Whites to be labeled a higher risk but not actually reoffend."

Judges were using misinformation derived from black box software to make life-changing decisions on whether someone is let free or receives a harsher sentence than appropriate.

Increasing the transparency of these programs and ensuring a diverse workforce is engaged on developing AI will help decrease bias and make software more inclusive. Increasing diversity among the AI workforce helps avoid the negative outcomes that can occur when AI development is concentrated among certain groups of individuals, including the risk of biases in AI systems.

As we move forward in this great age of technological modernization, I will be focused on how the private sector, Congress, and regulators can work together to ensure that AI technologies continue to innovate successfully and socially responsibly.

I want to thank our witnesses for testifying today and look forward to hearing your thoughts on how we can achieve this goal.

And, again, thank you, Mr. Chair.

Mr. HURD. I recognize the distinguished gentleman from Kentucky, Mr. Massie, is here. He is not a member of the subcommittee, so I ask unanimous consent that he is able to fully participate in this hearing. Without objection, so ordered.

Now I am pleased to announce and introduce our witnesses. Our first one, Dr. Amir Khosrowshahi, is vice president and chief technology officer of the Artificial Intelligence Products Group at Intel.

Welcome.

Dr. Charles Isbell is executive associate dean of the College of Computing within the Georgia Institute of Technology.

Dr. Oren Etzioni is the chief executive officer at the Allen Institute for Artificial Intelligence.

And Dr. Ian Buck is vice president and general manager of Accelerated Computing at NVIDIA.

Welcome to you all.

And pursuant to committee rules, all witnesses will be sworn in before you testify. So please rise and raise your right hand.

Do you solemnly swear or affirm that the testimony you are about to give is the truth, the whole truth, and nothing but the truth, so help you God?

Thank you.

Please let the record reflect that all witnesses answered in the affirmative.

In order to allow time for discussion, please limit your testimony to 5 minutes. Your entire written statement will be made part of the record.

And as a reminder, the clock in front of you shows your remaining time. The light will turn yellow when you have 30 seconds left, and when it turns red your time is up. And please remember to also push the button to turn on your microphone before speaking.

And now it is a pleasure to recognize Dr. Khosrowshahi for your initial 5 minutes.

WITNESS STATEMENTS

STATEMENT OF AMIR KHOSROWSHAHI

Mr. KHOSROWSHAHI. Good afternoon, Chairman Hurd, Ranking Member Kelly, and members of the House Committee on Oversight and Government Reform, Subcommittee on Information Technology.

My name is Amir Khosrowshahi, and I am the vice president and chief technology officer of Intel Corporation's Artificial Intelligence Products Group.

We're here today to discuss artificial intelligence, a term that was an aspirational concept until recently. While definitions of artificial intelligence vary, my work at Intel focuses on applying machine-learning algorithms to real world scenarios to offer benefits to people and organizations.

Thanks to technological advancements, AI is now emerging as a fixture in our daily lives. For instance, speech recognition features, recommendation engines, and bank fraud detection systems all utilize AI.

These features make our lives more convenient, but AI offers society so much more. For example, AI healthcare solutions will revolutionize patient diagnosis and treatment.

Heart disease kills one in four people in the United States. It is difficult for doctors to accurately diagnose disease, because different conditions present similar symptoms. That's why doctors mainly have had to rely on experience and instinct to make diagnoses. More experienced doctors tend to diagnose correctly three out of four times, those with less experience, however, just half the time, as accurate as the flipping of a coin. Patients suffer due to this information gap.

Recently, researchers using AI accurately spotted the difference between the two types of heart disease 9 out of 10 times. In this

regard, AI democratizes expert diagnoses for patients and doctors everywhere in the world.

AI is also contributing positively to agriculture. The population is growing, and by 2050 we will need to produce at least 50 percent more food to feed everyone. This will become increasingly challenging as societies will need to produce more food with less land to grow crops.

Thankfully, AI applications provide tools to improve crop yields and quality, while also reducing consumption of resources like water and fertilizer.

These are just a few examples of how AI is helping our communities. However, as we continue to harness the benefits of AI for societal good, governments will play a major role. We are in the early days of innovation of a technology that can do tremendous good. Governments should make certain to encourage this innovation and they should be wary of regulation that will stifle its growth.

At the Federal level, the United States Government can play an important role in enabling the further development of AI technology in a few ways.

First, since data fuels AI, the U.S. Government should embrace open data policies. To realize AI's benefits, researchers need to have access to large datasets. Some of the most comprehensive datasets are currently owned by the Federal Government. This data is a taxpayer-funded resource which, if made accessible to the public, could be utilized by researchers to train algorithms for future AI solutions.

The OPEN Government Data Act makes all nonsensitive U.S. Government data freely available and accessible to the public. Intel supports this bill and calls for its swift passage.

Second, the U.S. Government can help prepare an AI workforce. Supporting universal STEM education is a start, but Federal funding for basic scientific research at universities by agencies like the National Science Foundation is important to both train graduate-level scientists and contribute to our scientific knowledge base.

Current Federal funding levels are not keeping pace with the rest of the industrialized world. I encourage lawmakers to consider the tremendous returns on investment to our economy that funding science research produces.

In addition to developing the right talent to develop AI solutions, governments will have to confront labor displacement. AI's emergence will displace some workers, but too little is known about the types of jobs and industries that would be most affected.

Bills like H.R. 4829, the AI JOBS Act, help bridge that information gap by calling for the Labor Department to study the issue and to work with Congress on recommendations. Intel supports this bill as well and encourages Congress to consider it in committee.

AI promises many societal benefits, and government and industry should work together to harness them, and also to set up guidelines to encourage ethical deployment of AI and to prevent it from being used in improper ways that could harm the public.

I cannot stress enough how important it is that lawmakers seize the opportunity to enable AI innovation. As U.S. lawmakers con-

sider what to do in response to the emergence of AI, I encourage you to use a light touch. Legislating or regulating AI too heavily will only serve to disadvantage Americans, especially as governments around the world are pouring resources into tapping into AI's potential.

Thank you again for the opportunity to testify today. The government will play an important role in enabling us to harness AI's benefits while preparing society to participate in an AI-fueled economy. Determining whether or how existing legal and public policy frameworks may need to be altered will be an iterative process. Intel stands ready to be a resource as you consider these issues.

Thank you.

[Prepared statement of Mr. Khosrowshahi follows:]

Testimony of Dr. Amir Khosrowshahi, Vice President and Chief Technology Officer, Intel Corporation, Artificial Intelligence Products Group, before the U.S. House of Representatives, Committee on Oversight and Government Reform, Subcommittee on Information Technology, Hearing "Game Changers: Artificial Intelligence Part I"

February 14, 2018

Good afternoon, Chairman Hurd, Ranking Member Kelly, and members of the House Committee on Oversight and Government Reform, Subcommittee on Information Technology. My name is Amir Khosrowshahi, and I am the Vice President and Chief Technology Officer of Intel Corporation's Artificial Intelligence Products Group. Prior to joining Intel, I co-founded Nervana Systems, which created new hardware and software technology to allow companies to build Artificial Intelligence solutions. I have an academic background in neuroscience and machine learning from the University of California, Berkeley, and in mathematics and physics from Harvard University.

We are here today to discuss artificial intelligence. The term, "artificial intelligence" was first coined in the 1950's and until recently has been only an aspirational concept. Thanks to technological advances such as increased computing capability, large datasets, and innovations in algorithms, AI is now beginning to be a fixture in our daily lives. For instance, speech recognition, recommendation engines, and bank fraud detection systems are all features that utilize AI to make our lives more convenient. As CTO of Intel's AI Products Group, I can attest that despite this progress, the AI industry is still very much in the early days. However, as AI continues to advance, we must recognize that it will also cause shifts within technology, and within our society.

My work at Intel focuses on employing deep learning hardware chips and software algorithms to make inferences with data sets that can be applied to real-world scenarios offering benefits to organizations. In addition we spend a lot of time thinking about how AI can be harnessed to do good. There are several examples of how deep learning algorithms are currently being deployed to solve some of the world's most pressing challenges and benefitting society in turn.

AI solutions are poised to revolutionize the tools doctors employ to treat their patients. According to the Centers for Disease Control, about one in four deaths each year in the United States is caused by heart disease.[1] Detecting and treating various types of heart disease is tricky, and since doctors have long had to rely on experience and instinct in making diagnoses, patients were left with varying results.

Take for example the situation of a doctor treating a patient presenting common symptoms of heart failure. Making an accurate diagnosis is critical, and conditions like pericarditis and cardiomyopathy are often difficult to distinguish and therefore difficult to treat accurately. Even experienced doctors make the correct diagnosis only 3 out of 4 times. For doctors with less experience, their rate of correct diagnoses falls to 50 percent – as accurate as flipping a coin. In a recent experiment, researchers using AI spotted the difference between pericarditis and cardiomyopathy nine out of 10 times.[2] In this regard, artificial intelligence democratizes expert diagnoses for patients and doctors everywhere in the world.

---

[1] https://www.cdc.gov/heartdisease/facts.htm last accessed January 28, 2018

[2] https://www.intel.com/content/www/us/en/healthcare-it/article/improved-diagnosis.html (last accessed January 29, 2018)

AI is also contributing to agriculture improvements, providing tools to farmers to increase crop yield. The world's population will increase to 8.5 billion by the year 2030.[3] By 2050, we will need to produce at least 50 percent more food to feed the growing number of people in the world.[4] Meeting this demand will become increasingly difficult, for as populations climb, the supply of land decreases and societies must feed more people with less land to grow crops. Thankfully, with sensors, drones, robots, and advanced compute power, AI applications provide tools to farmers to deal with this problem. Intel technology helps farmers maximize crop yields, reduce environmental impact, and meet growing demand. "Precision agriculture" links sensor data on the ground to nearby Intel® IoT Gateways. The gateways act on that data locally and then send the collected data to cloud service providers, where they are analyzed using solutions from partners. These results should help farmers improve yields, reduce consumption of resources like water and fertilizer, and improve crop quality.

AI is also enabling law enforcement to combat child abuse. When Internet service providers identify instances of suspicious online activity pointing to child exploitation, they send these reports to the National Center for Missing and Exploited Children (NCMEC) - the nation's clearinghouse and comprehensive reporting center for all issues related to the prevention of and recovery from child victimization. From there, the NCMEC combs through these tips, sending legitimate leads to local law enforcement agencies. The problem is, NCMEC has a team of only 25 analysts employed to go through over 8 million tips. Intel partnered[5] with NCMEC to deliver an AI solution which helps analysts get the right information to the right jurisdiction. With this technology, NCMEC can scan sites for suspicious content, store massive volumes of data, run a variety of queries, and share the data across the organization's applications. AI helps automate and speed up the process, reducing the typical 30-day turnaround time to handle a report to just a day or two, which can save a child's life.[6]

As technologies evolve, they will undoubtedly continue to have an impact on our daily lives. Some have hailed the development of AI as evidence of the fourth industrial revolution. This analogy to the industrial revolution is telling – at that time through the turn of the nineteenth century, people were concerned with the same societal and economic shifts that are now being related to AI. Back then, some were concerned that they'd lose their jobs while others were hopeful for a 10 hour workweek.[7] In reality, while some industries – like horse-drawn carriages – faced severe displacement, others, like the automobile industry, flourished from the development of the combustion engine, creating many new

---

[3] http://www.un.org/sustainabledevelopment/blog/2015/07/un-projects-world-population-to-reach-8-5-billion-by-2030-driven-by-growth-in-developing-countries/ (last accessed January 29, 2018)

[4] https://www.intel.com/content/www/us/en/big-data/article/agriculture-harvests-big-data.html (last accessed January 29, 2018)

[5] https://www.intel.com/content/www/us/en/analytics/artificial-intelligence/article/ai-helps-find-kids.html (last accessed January 29, 2018)

[6] https://iq.intel.com/artificial-intelligence-is-good-for-society/ last accessed January 30, 2018

[7] Statement from Naveen Rao, Intel VP and GM of Intel's Artificial Intelligence Products Group at The Atlantic's event, The Innovation Game, November 7, 2017. Statement at approximately 26:10 of panel. Video available here: https://www.youtube.com/watch?v=aSDojU2HFWg&list=PLwj46yNDLyTV_V_Z5TrHMv1WE5WvK4QK-&index=2 last accessed February 6, 2018

jobs in the process. What didn't happen, despite that shift, was a 10 hour workweek. Indeed, with widespread improvements in productivity, there also tends to be creation of new jobs as well as development of new categories of jobs. As a society, our goal now should be figuring out how we can harness AI, tapping into the many benefits it holds, while mitigating the impacts of disruptions. And we need to embrace changes in technology and empower people to participate in an AI-fueled economy.

Governments will play a major role if we are to harness the benefits of AI while mitigating possible impacts. Broadly, as with any technological innovation, governments should focus their attention now on enabling the development of AI. We are in the early days of an innovation of a technology that can do tremendous good. Governments should make certain to encourage this innovation, and they should be wary of regulation that will stifle its growth. More specifically, Governments can assist in aiding the further development of this technology, ensuring citizens are equipped to participate in this emerging work economy, and safeguarding against potential disruptions to our way of life.

First, at the Federal level, the United States Government can play an important role in enabling the further development of AI technology. Since data is fuel for AI, the U.S. Government should embrace open data policies. For us to realize AI's benefits, researchers need to have access to large data sets. Some of the most comprehensive data sets are currently owned by the federal government. This data is a taxpayer-funded resource which, if made accessible to the public, could be utilized by researchers to train algorithms which would undergird future AI solutions. The OPEN Government Data Act makes all non-sensitive U.S. government data freely available and accessible to the public. Intel supports this bill and calls for its swift passage.

Second, the U.S. Government can help prepare an AI workforce. Workers need to have the right skills to create AI technologies and right now we have too few workers to do the job. Supporting universal STEM education is a start, but federal funding for basic scientific research that flows through agencies like the National Science Foundation to universities and helps to both train graduate level scientists and contribute to our scientific knowledge base on which industry can build are key. Current federal funding levels are not keeping pace with the rest of the industrialized world. I encourage lawmakers to consider the tremendous returns on investment to our economy that funding scientific research produces.

In addition to developing the right talent to develop AI solutions, governments will have to deal effectively with displaced workers. AI's emergence will displace some workers. But too little is known about the types of jobs and industries that will be affected most. Bills like H.R. 4829, the AI JOBS Act, that was recently introduced, help bridge that information gap by calling for the Department of Labor to study the issue and to work with Congress on recommendations. Intel supports this bill as well, and encourages Congress to consider it in Committee.

Finally, while AI holds many benefits, the government should play a role in helping society transition as AI emerges as a technology to mitigate undesirable impacts. Partnering with industry, government should set up guidelines to encourage ethical deployment of AI and to prevent it from being used in improper ways that would hurt the public. AI should not exist in "black boxes" – closed systems that receive an input, produce an output, and offer no explanation as to why or how[8] – especially as we

[8] https://ai.intel.com/the-challenges-and-opportunities-of-explainable-ai/ (last accessed January 29, 2018).

begin to rely on AI to help us with things like criminal sentencing, determining insurance policies, or the job candidate referral processes.

Ensuring transparent decisions, predictability, resistance to manipulation, and accountability will be important as we strive for "AI for good". Being able to explain why an algorithm made the decision it did will allow businesses and customers to make more informed choices about the AI tools that they use, and enable us to assign responsibility for outcomes to the proper human stakeholders. For these reasons, government should consider which AI uses require algorithm explainability to mitigate any potential for discrimination and harm to individuals.

As U.S. lawmakers consider whether and how – from legislative or regulatory standpoint – to deal with the emergence of AI, we encourage you to consider that technology like Artificial Intelligence which can enable innovation, economic growth and solve serious social issues, can only do so with a public policy environment that fosters its development. The United States has a long history of creating such an environment and to the extent it can continue to do so, it will reap the benefits of the technologies we are creating. Other countries see these potential benefits and are investing both in research and development and in creating favorable public policy environments. Intel recommends to all countries that they follow the proposals we include in our AI white paper, which I included as an addendum to my testimony. It is not a zero sum game and productive public policy environments in many countries can reinforce each other. The United States should specifically look at the proposals to make certain it does not fall behind.

Thank you again for the opportunity to testify today. We believe that AI is poised to create tremendous economic value while solving some of society's most pressing challenges, but governments must mitigate unwanted impacts. Striking the right balance will require governments to embrace changes and help society prepare. This will be an iterative process. As you consider how existing legal and policy frameworks may need to be altered to accomplish this, I hope you'll use Intel and other industry partners as a resource to offer technical expertise.

Thank you.

ADDENDUM

Artificial Intelligence

The Public Policy Opportunity

Intel and Artificial Intelligence

Intel powers the cloud and billions of smart, connected computing devices. Due to the decreasing cost of computing enabled by Moore's Law[9] and the increasing availability of connectivity, these connected devices are now generating millions of terabytes of data every day. Recent breakthroughs in computer and data science give us the ability to timely analyze and derive immense value from that data. As Intel distributes the computing capability of the data center across the entire global network, the impact of artificial intelligence is significantly increasing. Artificial intelligence is creating an opportunity to drive a new wave of economic progress while solving some of the world's most difficult problems. This is the artificial intelligence (AI) opportunity. To allow AI to realize its potential, governments need to create a public policy environment that fosters AI innovation, while also mitigating unintended societal consequences. This document presents Intel's AI public policy recommendations.

The **Collaborative Cancer Cloud**, a partnership between Intel, Oregon Health & Science University, Ontario Institute for Cancer Research and the Dana-Farber Cancer Institute, is enabling institutions to use distributed machine learning to speed up the discovery of new variants and biomarkers associated with cancer progression.

What is Artificial Intelligence?

While artificial intelligence is often equated with science fiction, it is not something looming on the horizon. It is already here, all around us, from the commonplace (talk-to-text, web searches, photo tagging, fraud detection) to the cutting edge (precision medicine, injury prediction, autonomous cars). Encompassing compute methods like advanced data analytics, computer vision, natural language

[9] According to Moore's Law, the number of transistors on a chip roughly doubles every two years. As a result, the scale gets smaller and transistor count increases at a regular pace to provide improvements in integrated circuit functionality and performance while simultaneously decreasing costs.

processing, semantic graphs, and machine and deep learning, AI is transforming the way businesses operate and how people engage with the world.

While there isn't a commonly accepted definition for AI, Intel views it as a computerized system that performs tasks we normally associate with people. But in spite of the remarkable advances of computing power and sophisticated algorithms, there is still a long way to go before what is called General AI becomes a reality. General AI refers to the ability of a computerized system to portray human-like intelligence across a multitude of tasks.

Intel is a founding partner of ***Hack Harassment***, a cooperative effort with the mission of reducing the prevalence and severity of online harassment. The initiative is evaluating AI technology and is working to develop an intelligent algorithm to detect and deter online harassment.

In contrast, Narrow AI, which addresses a specific task or set of tasks, is commonplace and its use in many sectors of society will only increase. These are the technologies that help scientists gain novel insights into cancer diagnosis and aid in the design of new therapies, help physicians identify risks and predict the onset of diseases, allow interacting with our phones or vehicles' navigation systems through speech recognition, power household cleaning robots, help financial institutions fight against fraud, assist us with driving our cars and make manufacturing safer and more productive. Many other examples exist and the important point is that AI will transform many aspects of our lives.

The Australian Government has funded an AI solution known as **Nadia**, an online virtual assistant who can speak, write, and chat online to interact with disabled persons. Nadia uses actress Cate Blanchett's voice and is programmed to improve its knowledge while answering thousands of questions asked in many different ways by people with diverse intellectual capabilities.

Before considering the public policy impact of AI, it is crucial to introduce "machine learning," an important technique behind many of the recent advances of AI. Machine Learning is the development and application of algorithms to build and continually improve models from data. It allows computers to "learn" without being explicitly programmed. It is particularly useful when it is hard for humans to explain their innate ability to infer one thing from another; for example, how to distinguish a cat from a dog? Or how to understand speech? It is also useful when the amount of data is enormous in relation to human reasoning abilities (in examples such as ranking web pages through a search engine). By providing large datasets as input (for example, millions of videos) the computer starts to recognize patterns such as images of cats, without anyone instructing the computer on what a cat looks like. Another example is machine learning's ability to analyze many spam email messages, thereby "learning" to differentiate normal mail from spam.

To take advantage of AI, all stakeholders must engage to understand the technology, debate how it will impact society and address concerns, as well as amplify its benefits and help society adjust.

Public Policy Considerations

The main drivers of public policy towards AI should be solving large societal problems and fostering economic progress. Accordingly, public policy must support industry efforts to bring AI benefits to the economy, to address citizens' concerns, and to identify needs for regulatory intervention.

As AI innovation is just beginning, it is crucial now to shape the public policy environment. Oversight by regulators will be essential for society to *trust* AI. Public policy should lower or remove any barriers standing between AI and its enormous potential to benefit our lives. Industry collaboration through groups like the Partnership on Artificial Intelligence[10] are important to further study the issues and develop specific solutions.

Intel proposes the following AI public policy principles, and offers specific recommendations for government implementation:

- **Foster Innovation and Open Development** – To better understand the impact of AI and explore the broad diversity of AI implementations, public policy should encourage investment in AI R&D. Governments should support the controlled testing of AI systems to help industry, academia, and other stakeholders improve the technology.

- **Create New Human Employment Opportunities and Protect People's Welfare** – AI will change the way people work. Public policy in support of adding skills to the workforce and promoting employment across different sectors should enhance employment opportunities while also protecting people's welfare.

- **Liberate Data Responsibly** – AI is powered by access to data. Machine learning algorithms improve by analyzing more data over time; data access is imperative to achieve more enhanced AI model development and training. Removing barriers to the access of data will help machine learning and deep learning reach their full potential.

[10] https://www.partnershiponai.org/#

- **Rethink Privacy –** Privacy approaches like The Fair Information Practice Principles and Privacy by Design have withstood the test of time and the evolution of new technology. But with innovation, we have had to "rethink" how we apply these models to new technology.

- **Require Accountability for Ethical Design and Implementation –** The social implications of computing have grown and will continue to expand as more people have access to implementations of AI. Public policy should work to identify and mitigate discrimination caused by the use of AI and encourage designing in protections against these harms.

Foster Innovation and Open Development

The potential of AI is enormous.[11] AI can enhance human capabilities, automate tedious or dangerous tasks keeping humans in safer conditions, unleash scientific discovery, and alleviate challenging societal problems. Doctors will be able to diagnose conditions earlier and more accurately, leading to quicker treatments and lives saved.[12] Automated vehicles will result in safer driving, and more efficiency and productivity. Farmers will increase crop yield based on real-time insights from weather and soil data, producing higher yields and more stable food supply even in unpredictable climates.

Realizing the potential of AI requires advances in core AI technologies. Governments must play a significant role in promoting those advances. Government investment in AI, public-private collaborations, and measures to incentivize adoption by society are public policy actions that will enable AI to develop and mature.

Equally, governments should gain expertise in AI in order to make effective public policy, to benefit from efficiency gains, and to champion AI adoption. Moreover, a new generation of AI specialists and data scientists should be on the radar of schools and universities when preparing new curricula.

Foster Innovation and Open Development – Recommendations

- **Fuel AI innovation:** Public policy should promote investment, make available funds for R&D, and address barriers to AI development and adoption.

[11] The global AI and robotics market is estimated to grow to $153 billion by 2020 (Robot revolution – Global robot and AI primer, Bank of America Merrill Lynch, Dec 2015.)

[12] The market for AI system in healthcare is estimated to grow from $633 million in 2014 to $6 billion in 2021 (From $600 M to $6 billion, AI systems poised for dramatic market expansion in healthcare, Frost & Sullivan, Jan 2016.)

- **Address global societal challenges:** AI-powered flagship initiatives should be funded to find solutions to the world's greatest challenges such as curing cancer, ensuring food security, controlling climate change, and achieving inclusive economic growth.
- **Allow for experimentation:** Governments should create the conditions necessary for the controlled testing and experimentation of AI in the real world, such as designating self-driving test sites in cities.
- **Prepare a workforce for AI:** Governments should create incentives for students to pursue courses of study that will allow them to create the next generation of AI.
- **Lead by example:** Governments should lead the way on demonstrating the applications of AI in its interactions with citizens and invest sufficiently in infrastructure to support and deliver AI-based services.
- **Partnering for AI**: Governments should partner with industry, academia, and other stakeholders for the promotion of AI and debate ways to maximize its benefits for the economy.

Create New Employment Opportunities and Protect People's Welfare

Productive work is a fundamental component of individual wellbeing and high functioning societies. In the same way that AI needs to be designed to function properly, so should society be prepared to leverage AI's benefits while mitigating its impact on the workforce. While AI has the potential to improve many aspects of our lives and to spur economic growth, AI and robotics will bring automation to broad categories of jobs (e.g. fully autonomous vehicles will reduce the need for trucking and taxi drivers). Concurrently, new tasks and jobs will be created requiring entirely different sets of skills. Governments need to understand how AI will impact employment and have a plan to encourage employment in ways that allow technology to assist humans in the pursuit of their work.

From more timely, more accurate medical diagnostics to intelligent, safer transportation, AI will affect all facets of the economy, including the public sector. The economic benefits of AI should be inclusive, accessible, and broadly shared by society. Public policies must be enabled to mitigate inequalities, protect citizens' welfare, and help with the transition to a more data-driven economy.

Create New Employment Opportunities and Protect People's Welfare – Recommendations

- **Encouraging Human Employment:** Governments should implement programs to mitigate AI's impact on jobs and devise policies that promote employment. These programs should particularly focus on the effectiveness of incentives in government funded infrastructure projects.
- **Retraining:** Governments should implement policies that support the up-skilling and the re-skilling of the workforce, particularly in job areas that are less likely to be automated, such as positions focused on person to person interaction and the need for "guided computation" where individuals direct and oversee the operation of the technology.

Liberate Data Responsibly

AI does not exist without data. Machine learning based algorithms are trained with existing data and those data relate to specific usage domains. For instance, if AI is to be used to fight cancer, then deidentified data from medical records, genomic information, state of the art treatments and many other domains should be made available. Of particular interest are solutions that allow for the federated access to data from distributed repositories held in different sites, while preserving privacy and security.

Governments are also solicitors, creators and repositories of data. As long as no personal or sensitive information is involved, many of these datasets should be made available for public use. If personal or sensitive information is a requirement to solve critical societal problems (like making breakthroughs in personalized medicine), governments should partner with the private sector to find solutions to use AI while still delivering privacy protections. One example of such protection is the use of AI algorithms that analyze data in several encrypted yet separate datasets, but never require sharing of the data outside the encrypted area. These mechanisms for "raw data", instead of aggregated inferences, are much more useful for training data analytics necessary for AI. As explained before, AI requires data to function and public sector data is a valuable source of information to develop AI solutions to societal challenges.

Liberate Data Responsibly – Recommendations

- **Keep data moving:** Governments should eliminate unwarranted data localization mandates and enable secure international data transfers through international agreements and legal tools.
- **Open public data:** While protecting privacy, governments should make useful datasets publicly available when appropriate and provide guidance to startups and small and medium businesses for its reuse.
- **Support the creation of reliable data sets to test algorithms:** Governments should explore non-regulatory methods to encourage the development of testing data sets.
- **Federate access to data**[13]**:** Governments should partner with industry to promote AI tools to access encrypted data for analysis, while not requiring transfer of the data.

Rethink Privacy

Where the data used for AI originates from identifiable individuals, appropriate protections should be implemented to ensure that data is deidentified, lawfully accessed, processed, and kept safe. Robust privacy regulatory frameworks for the protection of personal data and cybersecurity should also apply to AI implementations. Intel is a proponent of technology neutral comprehensive privacy laws based on the Organization for Economic Cooperation and Development's Fair Information Practice Principles (the FIPPs), which are the global common language of privacy.

Intel has long supported and implemented Privacy by Design.[14] Intel has previously demonstrated through our Rethinking Privacy[15] project, that the FIPPs can be implemented during Privacy by Design processes to better protect individuals.

Questions may arise regarding the enforceability of privacy protections when a machine uses data autonomously. In these circumstances, accounting for privacy principles when designing technology will help protect individuals.

"Security Safeguards" is one of the FIPPs and is particularly critical in protecting the trustworthiness of AI implementations. AI can be used to foster both privacy and security by predicting the spread of cybersecurity attacks and helping organizations protect their data and

[13] Instead of centralizing data from several institutions, federated access to data allows each institution to keep control of their data while enabling joint data analytics across all institutions.
[14] *Privacy by Design* refers to the philosophy and approach of embedding privacy into the design specifications of various technologies.
[15] http://blogs.intel.com/policy/files/2015/01/RethinkingPrivacy.pdf

AI algorithms/models. A critical component of allowing AI to better protect privacy and security will be the use of cybersecurity data to better predict future attacks. As the compute power of the data center is distributed across the entire network, the potential for AI to stop cyberattacks before they do significant harm will be greatly increased. This is one of many reasons why governments should promote the use and sharing of data for cybersecurity purposes.

Rethink Privacy – Recommendations

- **Adopt Robust Privacy Laws:** Based on the OECD Fair Information Practice Principles.
- **Implement Privacy by Design:** Follow Intel's Rethinking Privacy approach to implement Privacy by Design into AI product and project development.
- **Keep data secure:** Policies should help enable cutting-edge AI technology with robust cyber and physical security to mitigate risks of attacks and promote trust from society.
- **It takes data for AI to protect data:** Governments should adopt policies to reduce barriers to the sharing of data for cybersecurity purposes.

Require Accountability for Ethical Design and Implementation

Trust in AI requires organizations to demonstrate to the public and government regulating bodies that the technology is designed, implemented, and operated responsibly.

The Information Accountability Foundation (IAF [16]) has spent considerable time articulating the essential elements of what is required to demonstrate the responsible handling of information. The IAF's five principles are:

1. Organization commitment to accountability and adoption of internal policies consistent with external criteria.
2. Mechanisms to put privacy policies into effect, including tools, training and education.
3. Systems for internal ongoing oversight and assurance reviews and external verification.
4. Transparency and mechanisms for individual participation.
5. Means for remediation and external enforcement.

With only small adjustments (amending the word "privacy" in the second principle to cover broader categories of automated decision making), this work can and should apply more broadly to AI. Organizations that develop and implement AI solutions will benefit from working through the principles as the resulting policies, processes, and resources put in place will demonstrate responsible behavior to both regulators and individuals who are impacted by AI solutions.

[16] http://informationaccountability.org/

Applying the principles to AI requires new thinking. As an example, transparency may be more difficult for some AI approaches than with traditional data processing. Some algorithms use hundreds of millions of adjustable parameters to function and may be continually updated based upon real-time data. In some cases this makes it impossible to deconstruct how a particular result was produced by the algorithm to accurately trace back a cause. In other words, it may be impossible to understand how a result is achieved, consequently making AI less accountable to the user. However, there is ongoing research to derive rules from deep neural networks, and these algorithms are being used successfully, for example for sensory recognition (like image recognition and natural language speech interfaces) and fraud detection by financial institutions. There are also new approaches to AI such as memory-based reasoning that can better warrant outputs, including reference back to relevant prior episodes of similar persons, places, and things as raw evidence in the explanations to the user. More symbolic approaches to AI also claim transparency in backtracking of inferential logic chains.

Ensuring fairness of AI results depends upon how the algorithms were developed, and in the case of AI-based machine learning, also upon the data that was utilized for their training. Noting that AI algorithms have the potential to make less biased decisions than people, there is still a risk for unintended bias, and therefore unintended discrimination of individuals. This may happen, for example, when the data used to train the algorithm was not representative of the problem space in question. One example of this situation could be when the training datasets were not free from bias themselves. Means to mitigate bias include using algorithms and data models that account for bias, well-curated training sets, extensive verification and validation of AI systems, and alertness to possible ethical or fairness implications from AI-based decisions. Government and the private sector should continue to work together to study and develop solutions to regulate discrimination caused by AI implementations.

Require Accountability for Ethical Design and Implementation – Recommendations

- **Standing for "Accountable Artificial Intelligence":** Governments, industry and academia should apply the Information Accountability Foundation's principles to AI. Organizations implementing AI solutions should be able to demonstrate to regulators that they have the right processes, policies and resources in place to meet those principles.
- **Transparent decisions:** Governments should determine which AI implementations require algorithm explainability to mitigate discrimination and harm to individuals.

SUMMARY OF RECOMMENDATIONS

FOSTER INNOVATION AND OPEN DEVELOPMENT

- **Fuel AI innovation**: Public policy should promote investment, make available funds for R&D, and address barriers to AI development and adoption.
- **Address global societal challenges:** AI-powered flagship initiatives should be funded to find solutions to the world's greatest challenges such as curing cancer, ensuring food security, controlling climate change, and achieving inclusive economic growth.
- **Allow for experimentation:** Governments should create the conditions necessary for the controlled testing and experimentation of AI in the real world, such as designating self-driving test sites in cities.
- **Prepare a workforce for AI:** Governments should create incentives for students to pursue courses of study that will allow them to create the next generation of AI.
- **Lead by example:** Governments should lead the way on demonstrating the applications of AI in its interactions with citizens and invest sufficiently in infrastructure to support and deliver AI-based services.
- **Partnering for AI**: Governments should partner with industry, academia, and other stakeholders for the promotion of AI and debate ways to maximize its benefits for the economy.

CREATE NEW EMPLOYMENT OPPORTUNITIES AND PROTECT PEOPLE'S WELFARE

- **Encouraging Human Employment:** Governments should implement programs to mitigate AI's impact on jobs and devise policies that promote employment. These programs should particularly focus on the effectiveness of incentives in government funded infrastructure projects.
- **Retraining:** Governments should implement policies that support the up-skilling and the re-skilling of the workforce, particularly in job areas that are less likely to be automated, such as positions focused on person to person interaction and the need for "guided computation" where individuals direct and oversee the operation of the technology.

LIBERATE DATA RESPONSIBLY

- **Keep data moving:** Governments should eliminate unwarranted data localization mandates and enable secure international data transfers through international agreements and legal tools.
- **Open public data:** While protecting privacy, governments should make useful datasets publicly available when appropriate and provide guidance to startups and small and medium businesses for its reuse.
- **Support the creation of reliable data sets to test algorithms:** Governments should explore non-regulatory methods to encourage the development of testing data sets.
- **Federate access to data**[17]**:** Governments should partner with industry to promote AI tools to access encrypted data for analysis, while not requiring transfer of the data.

RETHINK PRIVACY

- **Adopt Robust Privacy Laws:** Based on the OECD Fair Information Practice Principles.
- **Implement Privacy by Design:** Follow Intel's Rethinking Privacy approach to implement Privacy by Design into AI product and project development.
- **Keep data secure:** Policies should help enable cutting-edge AI technology with robust cyber and physical security to mitigate risks of attacks and promote trust from society.
- **It takes data for AI to protect data:** Governments should adopt policies to reduce barriers to the sharing of data for cybersecurity purposes.

REQUIRE ACCOUNTABILITY FOR ETHICAL DESIGN AND IMPLEMENTATION

- **Standing for "Accountable Artificial Intelligence":** Governments, industry and academia should apply the Information Accountability Foundation's principles to AI. Organizations implementing AI solutions should be able to demonstrate to regulators that they have the right processes, policies and resources in place to meet those principles.
- **Transparent decisions:** Governments should determine which AI implementations require algorithm explainability to mitigate discrimination and harm to individuals.

[17] Instead of centralizing data from several institutions, federated access to data allows each institution to keep control of their data while enabling joint data analytics across all institutions.

Mr. HURD. Thank you, Dr. Khosrowshahi.

Dr. Isbell, you are now recognized for 5 minutes.

STATEMENT OF CHARLES ISBELL

Mr. ISBELL. Chairman Hurd, Ranking Member Kelly, and distinguished members of the subcommittee, my name is Dr. Charles Isbell. I am a professor and executive associate dean for the College of Computing at Georgia Tech. I would like to thank you for the opportunity to appear before the subcommittee.

As requested by the subcommittee, my testimony today will focus on the potential for artificial intelligence and machine learning to transform the world around us and how we might collectively best respond to this potential.

There are many definitions of AI. My favorite one is that it is the art and science of making computers act the way they do in the movies. In the movies, computers are often semimagical and anthropomorphic. They do things that if humans did them, we would say they required intelligence.

As noted by the chairman, if that is AI, then we already see AI in our everyday lives. We use the infrastructure of AI to search more documents than any human could possibly read in a lifetime, to find the answers to a staggering variety of questions, often expressed literally as questions. We use that same infrastructure to plan optimal routes for trips, even altering our routes on the fly in the face of changes in traffic.

We let computers finish our sentences, sometimes facilitating a subtle shift from prediction of our behavior to influence over our behavior. And we take advantage of these services by using computers on our phones or home speakers to interpret a wide variety of spoken commands.

All of this is made possible because AI systems are fundamentally about computing and computing methods for automated understanding and reasoning, especially ones that leverage data to adapt their behavior over time.

That AI is really computing is an important point to understand. What has enabled many of the advances in AI is the stunning increase of computational power, combined with the ubiquity of that computing.

That AI also leverages data is equally important. The same advances in AI are also due, in large part, to the even more stunning increase in the availability of data, again made possible by ubiquity, in this case of the internet, social media, and relatively inexpensive sensors, including cameras, GPS, microphones, all embedded in devices we carry with us, connected to computers that are, in turn, connected to one another.

By leveraging computing and data, we are moving from robots that assemble our cars to cars that almost drive themselves. One can be skeptical, as I am, that we will in the near future create AI that is as capable as humans are in performing a wide variety of the sort of general tasks that humans grapple with every day simultaneously. But it does seem that we are making strong progress toward being able to solve a lot of very hard individual tasks as well as humans.

We may not replace all 3 million truck drivers and taxi cab drivers, nor all 3 million cashiers in the United States, but we will increasingly replace many of them. We may soon trust the x-ray machine itself to tell us whether we have a tumor as much as we trust the doctor. We may not automate away intelligence analysts, but AI will shape and change their analysis.

So AI exists and is getting better. It is not the AI of science fiction, neither benevolent intelligence working with humans as we traverse the galaxy, nor malevolent AI that seeks humanity's destruction. Nonetheless, we are living every day with machines that make decisions that if humans made them we would attribute to intelligence.

As noted by the ranking member, it is worth noting that these machines are making decisions for humans and with humans. Many AI researchers and practitioners are engaged in what we might call interactive AI. The fundamental goal there is to understand how to build intelligent agents that must live and interact with large numbers of other intelligent agents, some of whom may be human.

Progress towards this goal means that we can build artificial systems that work with humans to accomplish tasks more effectively, can respond more robustly to changes in the environment, and can better coexist with humans as long-lived partners.

But as with any partner, it is important that we understand what our partner is doing and why. To make the most of this emerging technology, we will need a more informed citizenry, something we can accomplish by requiring that our AI partners are more transparent on the one hand and that we are more savvy on the other.

By transparency, I mean something relatively simple. An AI algorithm should be inspectable. The kind of data the algorithm uses to build its model should be available. It is useful to know that your medical AI was trained to detect heart attacks mostly in men.

The decisions that the system makes should be explainable and understandable. In other words, as we deploy these algorithms, each algorithm should be able to explain its output and its decisions: This applicant was assigned higher risk because is not only more useful, but is less prone to abuse than just this applicant was assigned a higher risk.

To understand such machines, much less to create them, we have to strive for everyone to not only be literate, but to be compurate. That is, they must understand computing and computational thinking and how it fits into problem-solving in their everyday lives.

I am excited by these hearings. Advances in AI are central to our economic and social future. The issues that are being raised here are addressable and can be managed with thoughtful support for robust funding and basic research in artificial intelligence, as noted by my colleague, support for ubiquitous and equitable computing education throughout the pipeline, in K–12 and beyond, and the developing standards for the proper use of intelligent systems.

I thank you very much for your time and attention today, and I look forward to working with you in your efforts to understand how we can best develop these technologies to create a future where we are partners with intelligent machines.

Thank you.
[Prepared statement of Mr. Isbell follows:]

House Committee on Oversight and Government Reform
Subcommittee on Information Technology
February 14, 2018

"Game Changers: Artificial Intelligence Part I"

Witness: Charles Isbell, Georgia Institute of Technology

Chairman Hurd, Ranking Member Kelly, and distinguished members of the subcommittee, my name is Dr. Charles Isbell and I am a Professor and Executive Associate Dean for the College of Computing at Georgia Tech. Thank you for the opportunity to appear before this Subcommittee to discuss the development, uses, barriers to adoption, and potential challenges and advantages of government use of artificial intelligence.

By way of explaining my background, let me note that while I tend to focus on statistical machine learning, my research passion is actually artificial intelligence. I like to build large integrated systems, so I also tend to spend a great deal of my time doing research on autonomous agents, interactive entertainment, some aspects of human-computer interaction, software engineering, and even programming languages

I think of my field as interactive artificial intelligence. My fundamental research goal is to understand how to build autonomous agents that must live and interact with large numbers of other intelligent agents, some of whom may be human. Progress towards this goal means that we can build artificial systems that work with humans to accomplish tasks more effectively; can respond more robustly to changes in environment, relationships, and goals; and can better co-exist with humans as long-lived partners.

As requested by the Subcommittee, my testimony today will focus on the potential for artificial intelligence and machine learning to transform the world around us. I will:

1. Situate recent developments in AI in the larger context of developments in computing more generally;
2. Explore the potential uses and misuses of this technology by focusing on the human-machine loop; and
3. Discuss the gaps in education and training that threaten to minimize participation in the field.

As the members of this Subcommittee well know, there has been an explosion in the development and deployment of what we might call AI technology. With that explosion has come a corresponding explosion in interest in AI.

In any discussion—particularly technical ones—it helps to define our terms. There are many potential definitions of AI. My favorite one is that it is "the art and science of making computers act like they do in the movies." In the movies, computers are often semi-magical and anthropomorphic; they do things that, if humans did them, we would say they required intelligence. This definition is borne out in our use of AI in the everyday world. We use the infrastructure of AI to search billions upon billions of documents to find the answers to a staggering variety of questions—often expressed literally as questions. We use automatically tagged images to organize our photos, and we use that same infrastructure to plan optimal routes for trips—even altering our routes on-the-fly in the face of changes in traffic. We are able to detect automatically the presence of tumors from x-rays, even those trained doctors find difficult to see. We let computers finish our sentences as we type texts and use search engines, sometimes facilitating a subtle shift from prediction of our behavior to influence over our behavior. Often we take advantage of these services by using our phones (our phones!) to interpret a wide variety of spoken commands.

This basic definition, of course, ignores what is going on underneath the hood. Perhaps a somewhat better way of grappling with AI is to understand that it is a set of computing techniques and approaches that attempt to solve exponentially hard problems in reasonable time by cheating in clever ways. In other words, at bottom, AI is about *computing methods* for automated understanding and reasoning, especially ones that *leverage data* to adapt their behavior over time.

That AI is really about computing is important. What has enabled many of the advances in AI over the last decade is the stunning increase of computational power combined with the ubiquity of that computing. That AI also leverages data is equally important. The advances in AI over the last decade are also due in large part to the even more stunning increase in the availability of data, again made possible by the ubiquity of the internet, social media, and relatively inexpensive sensors (including cameras, GPS, and the computer itself) that track our every move.

Consider the problems above: Google leverages arrays of server farms to index and search documents now available digitally; neural networks use millions of examples of pictures of human faces to perform the hundreds of millions of calculations necessary to do face-tagging in the same way that we can do phoneme and word detection from audio; our navigation apps like Waze use both the digital expression of maps to sort through millions or even billions of possible paths from one place to another, as well as the ubiquity of GPS in other vehicles to detect changes in traffic; medical prediction software can build tumor detectors by leveraging decades of data on x-rays and ground-truth labels of cancer; and the same techniques are used to crowd-source likely completions to search queries.

Consider the technology behind them: Deep learning is an update on an algorithm whose modern expression was known about the time of my birth. It uses layers of

artificial "neurons" to map from a set of features (*e.g.*, pixels, sounds, financial information, and so on) to more abstract concepts (*e.g.*, names of objects, words, credit-worthiness, and so on). As recently as twenty years ago, computational power was such that one could only build one or two layers. Performance required highly trained humans hand-tuning both network structure and the form of features themselves. Now, with both cheap, fast computing power and an abundance of data, the structure and features can also be learned, freeing computing professionals to develop better techniques that take advantage of this newfound power. Accordingly, the new systems work far better than we had available even a few years ago.

So, in some very important sense, AI already exists. It is not the AI of science fiction, neither benevolent intelligences working with humans as we traverse the galaxy, nor malevolent AI that seeks humanity's destruction. Nonetheless, we are living every day with machines that make decisions that, if humans made them, we would attribute to intelligence. And the machines often make those decisions faster and better than humans would.

Importantly, each of the examples we consider above is a distinctly human-centered problem. It is human-centered both in the sense that these systems are trying to solve problems that humans deal with every day—question answering, symptom evaluation, navigation—but also human-centered in the sense that humans have or currently perform some of those tasks. Presumably, these developments are all to the good. We are living up to the promise of technology that allows us to automate away work that is dirty, dangerous, or dull, freeing up human capital to be more productive, and, hopefully, for humans to be more fulfilled. The social and economic benefits are potentially immense.

There are also some reasons for concern. Those concerns also have social and economic components. I will focus briefly on two potential issues: the opaqueness of our intelligent machines and the potential impact on jobs.

We are increasingly using our AI systems to make real decisions, and we do not necessarily understand those decisions. In particular, these decisions can have severe impacts. For example, according to the Marshall Project, almost every state uses some form of "risk assessment" at some stage in the criminal justice system.

> Risk assessments have existed in various forms for a century, but over the past two decades, they have spread through the American justice system, driven by advances in social science. The tools try to predict recidivism — repeat offending or breaking the rules of probation or parole — using statistical probabilities based on factors such as age, employment history, and prior criminal record. They are now used at some stage of the criminal justice process in nearly every state. Many court systems use the tools to guide decisions about which prisoners to release on parole, for example, and risk as-

sessments are becoming increasingly popular as a way to help set bail for inmates awaiting trial.

Consider the automation of this process, relying on an algorithm in lieu of a judge's discretion. As noted by Cathy O'Neil, author of *Weapons of Math Destruction*, the data used by these algorithms to build models are sometimes suspect. Worse, we treat the output as "objective" without understanding that the data are themselves not objective. In this particular case, we set out to predict recidivism as if that means *the chance of committing a crime again* when in fact we are predicting *the chance of being arrested and convicted again.* It does not take much imagination to see how being from a heavily policed area raises the chances of being arrested again, being convicted again, and in aggregate leads to even more policing of the same areas, creating a feedback loop. One can imagine similar issues with determining fit for a job, or credit-worthiness, or even face recognition and automated driving. In computing, we call this garbage-in-garbage-out: an algorithm is only as good as its data. This saying is certainly true, and especially relevant for AI algorithms that learn based on the data they are given.

Luckily, one way to address these issues is straightforward: to increase transparency. An AI algorithm should inspectable. The kind of data the algorithm uses to build its model should be available. The decisions that such algorithms make should be inspectable. In other words, as we deploy these algorithms, each algorithm should be able to explain its output. "This applicant was assigned high risk because..." is more useful than, "This applicant was assigned high risk."

Of course, as we make our AI better and easier to understand, it is difficult not to imagine that AI will do more and more for us. In today's climate, we are imagining not only robots that assemble our cars, but that those cars will drive themselves. We can see a world where we will not only have algorithms that allow us to watch the stock market but will do a faster, better job buying and selling stocks than stockbrokers do. We may soon trust the x-ray machine itself to tell us if we have a tumor more than we trust a doctor. I am skeptical that we will create such AI machines in the near future, but it does seem that we are making inexorable progress toward that end. We may not replace all truck drivers and taxi cab drivers, but we may replace many of them. We may not replace all cashiers, but we may replace many of them. In a country where there are nearly 3 million truck drivers and more than 3 million cashiers, one can imagine what a significant impact such automation will have on the economy and on the job force.

Luckily again technology and automation does not simply destroy jobs, it creates them. In this particular case, it creates jobs that require technological sophistication and understanding. Here, it is important to return to our definitions. AI is about computing methods for automated understanding and reasoning, especially ones that leverage data to adapt their behavior over time. Thus, the future belongs to those who are not simply highly literate but compurate; that is, those who under-

stand computing and how it fits into problem solving will be most productive and impactful in the future.

We can see in the current data that our fellow citizens understand this reality. At Georgia Tech, we launched an affordable online master's degree in Computer Science four years ago. We are currently enrolling 6,365 students, 70% of whom are US citizens or permanent residents. Across the country, undergraduate computer science enrollments are at an all-time high at Research I universities, growing 113% between 2009 and 2015. From 2006 to 2015, the average number of CS majors increased for large departments (10+ faculty) from 320 to 970 and for small departments from 160 to 500 majors. The overall numbers are significantly higher than at the height of the dot-com boom.

At the same time, non-majors are increasingly taking upper-division computing courses for use in their own fields. According to Generation CS, the number of non-majors in courses intended for majors is increasing at a rate equal to or higher than that of majors. We are also seeing increasing interest in AI. For example, at Georgia Tech, 43% of our CS minors are focused on Artificial Intelligence. This year, our peers are reporting record numbers of graduate student applicants in machine learning and artificial intelligence.

Even more telling, institutions have been forced to cap the number of students who major in a program. This throttling of support suggests that demand may be even higher than it seems, but it also suggests that we are not capable of responding to this demand even as we need to educate more and more students in the area. The number of Ph.D. graduates in computer science going into higher education is dropping significantly. Further, this issue is not limited to those seeking undergraduate and advanced degrees. We are seeing an increasing need to educate students at the high school level as well and a corresponding lack of teachers available who are qualified to teach foundational computer science in K-12. Given the slow pace of production and the lack of an incentive structure for graduates in computer science to become teachers, the country will not be able to produce enough CS teachers quickly enough to meet demand.

In Georgia, for example, there are approximately 519,000 high school students. Only 29,000 of them are enrolled in computing of any kind—less than 6 percent. According to the Professional Standards Commission, the governing body over teacher certification in the state, there were only 93 credentialed teachers in 2017. The majority of the computing courses in the state are being taught by approximately 400-500 engaged and committed teachers who are not certified to do so. The state is in its nascent stages of offering a framework to guide what "high school-level CS" actually means. For now, the curricula and quality of the CS courses vary tremendously. The College Board's Advanced Placement Computer Science A exam is more formalized and demonstrates the magnitude of the problem for rigorous computing. Data from the College Board suggest that, in 2017, only 125 of the 500 high schools in the state

offered AP Computer Science. In Atlanta Public Schools, which is in the heart of Georgia's technology hub, there are only two high schools that offer Advanced Placement Computer Science.

Under these circumstances, possibly the only way to deploy this subject broadly is to offer blended learning courses. The core content of computational courses will ultimately have to be delivered through online platforms in close conjunction with classroom teachers who can be present and facilitate the actual process of learning.

In conclusion, I am excited by these hearings. Advances in AI are central to our economic and social future. The issues are being raised here can be addressed with thoughtful support for robust funding in basic research in artificial intelligence—including research in how to engage in education; support for that education throughout the pipeline; and in developing standards for the proper use of intelligent systems.

I thank you very much for your time and attention today, and I look forward to working with you in your efforts to understand how we can best develop these technologies to create a future where we are partners with intelligent machines.

Thank you. This concludes my testimony.

Mr. HURD. Thank you, sir.

Dr. Etzioni, you are now up for 5 minutes.

STATEMENT OF OREN ETZIONI

Mr. ETZIONI. Good afternoon, Chairman Hurd and Ranking Member Kelly, distinguished members of the committee. Thank you for the opportunity to speak with you today about the nature of AI and the role of the Federal Government.

My name is Oren Etzioni. I am the CEO of the Allen Institute for Artificial Intelligence, which is backed by Paul Allen. We call ourselves AI2. Founded in 2014, AI2 is a nonprofit research institute whose mission is to improve lives by conducting high-impact research and engineering in the field of AI for the common good.

The goal of my brief remarks today is to help demystify AI and cut through a lot of the hype on the subject. And I'm delighted to talk to you in particular, Chairman, with a computer science degree. But it's really important to me to make sure that my remarks are understandable by everybody and that we don't confuse science fiction with the real science and Hollywood and hype with what's actually going on.

What we do have are these very narrow systems that are increasingly sophisticated, but they're also extremely difficult to build. We need to work to increase the supply of people who can do this. And that's going to be achieved through increased diversity, but also through immigration.

And so, so many of us are immigrants to this country. At AI2, we have 80 people who come from literally all over the world, from Iran, from Israel, from India, et cetera, et cetera. We need to continue to welcome these people so we can continue to build these systems.

I have a number of thoughts, but I actually want to address the issue that came up just in the conversation now about transparency and bias and certainly the concerns that we have about these database systems generating unfairness.

Obviously, we want the systems to be fair, and obviously, we want them to be transparent. Unfortunately, it's not as easy as it sounds. These are complex statistical models that are ingesting enormous amounts of data, millions and billions of examples, and generating conclusions.

So we have to be careful. And I think the phrase "light touch" is a great one here. We have to be very careful that we don't legislate transparency, but rather that we attempt to build algorithms that are more favored, more desired, because they're more transparent.

I think legislating transparency or trying to do that would actually be a mistake, because ultimately consider the following dilemma. Let's say you have a diagnostic system that's highly transparent and 80 percent accurate. You've got another diagnostic system that's making a decision about a key treatment. It's not as transparent, okay, that's very disturbing, but it's 99 percent accurate. Which system would you want to have diagnosing you or your child?

That's a real dilemma. So I think we need to balance these issues and be careful not to rush to legislate what's complex technology here.

While I'm talking about legislation and regulation and the kinds of decisions you'll be making, I want to emphasize that I believe that we should not be regulating and legislating about AI as a field. It's amorphous. It's fast-moving. Where does software stop and AI begin? Is Google an AI system? It's really quite complicated.

Instead, I would argue we should be thinking about AI applications. Let's say self-driving cars. That's something that we should be regulating, if only because there's a patchwork of municipal and State regulations that are going to be very confusing and disjointed, and that's a great role for the Federal Government.

The same with AI toys. If Barbie has a chip in it and it's talking to my child, I want to be assured that there are some guidelines and some regulations about what information Barbie can take from my child and share publicly. So I think that if we think about applications, that's a great role for regulation.

And then the last point I want to make is that we need to remember that AI is a tool. It's not something that's going to take over. It's not something that's going to make decisions for us, even in the context of criminal justice. It's a tool that's working side by side with a human.

And so long as we don't just rubber stamp its decisions but rather listen to what it has to say but make our own decisions and realize that maybe AI ought to be thought of as augmented intelligence rather than artificial intelligence, then I think we're going to be in great shape.

Thank you very much.

[Prepared statement of Mr. Etzioni follows:]

Written Testimony of
Oren Etzioni
CEO, Allen Institute for Artificial Intelligence
Professor of Computer Science
University of Washington

Hearing on:
Game Changers: Artificial Intelligence Part I

before the United States the
House Committee on Oversight and Government Reform
Subcommittee on
Information Technology

February 14, 2018

Good morning, Chairman Hurd and Ranking Member Kelly, and distinguished Members of the Committee. Thank you for the opportunity to speak with you today about the nature of Artificial Intelligence and role of the federal government. My name is Oren Etzioni, and I am the CEO of the Allen Institute for Artificial Intelligence (AI2). Founded in 2014, AI2 is a non-profit research institute whose mission is to improve lives by conducting high-impact research and engineering in the field of AI, all for the common good.

The goal of my brief remarks today is to help demystify AI, and to cut through a lot of the hype on the subject.

- What is AI? Let's not confuse science with science fiction
- Superintelligence is over-hyped—we are building narrow systems
- Distinguish intelligence and autonomy
- Regarding regulation:
    1. We can't afford to stifle innovation due to global competition
    2. Regulating AI itself is challenging as it's fast moving and is difficult to distinguish from software
    3. We can and should regulate AI applications---if only to prevent a patchwork of state rules as is the case with self-driving cars
- Principles:
    1. AI should have an off switch
    2. AI should associate clear responsibility (if not liability) with a person—it's a tool not a mask to hide behind---"my AI did it is not an excuse"
    3. AI should identify itself --- think of the case of bots, of fake porn
    4. AI should respect privacy--- think of AI barbie talking to our kids
    5. AI should not exacerbate bias in data. Data is historical, and yet AI is predictive

Let me start by setting things up. Elon Musk, the tech entrepreneur, has told us that with AI we're summoning the demon, that AI represents an existential threat to the human race. We see headlines in the newspapers all the time. Here's one from *Newsweek*: "Artificial intelligence is coming and it could wipe us out." Is AI really poised to wipe us out? The famous roboticist, Rod Brooks, said if you're worried about the terminator, just close your door. Perhaps these robots are not as threatening as they're made out to be.

Andrew Ng from Stanford, founder of Coursera, said that working to prevent AI from turning evil is like disrupting the space program to prevent overpopulation on Mars. It ignores the technical difficulties. It also ignores the potential benefits. Stephen Hawking, in recent remarks, said AI is either the best thing that will happen to the human race or the worst thing. It's nowhere in the middle. It's one or the other.

Before I continue, I'd love to get a sense from you. How many people are really worried about AI and its impact? How many people see it as a major positive for our society? Okay, it's a techy crowd. Maybe I should quit while I'm ahead. I don't know if you saw, a lot of hands went up on the positive side.

Let's talk about this in more depth. I think the first thing we need to do is really separate science from science fiction and the Hollywood hype from the realities of AI. We need to ask ourselves, 'What is AI today?' The answer to that question is certainly very different than what you see in Hollywood films, what you see in "Westworld," etc.

What is AI? First of all, for many years AI was a lot of hype. That's something that's changed. We have seen a number of very real AI successes. We've seen an AI program defeat the world champion in chess. We've seen AI do speech recognition very well. That's the basis of Siri, Alexa and systems like that. We have systems that can now do facial recognition with very high accuracy. That's used in security. That's used by Facebook. It's used in a variety of ways.

"We've seen an AI program defeat the world champion in chess. We've seen AI do speech recognition very well. That's the basis of Siri, Alexa and systems like that. We have systems that can now do facial recognition with very high accuracy."

We have very powerful machine translation, for those of you that have used translate.google.com or other such apps. Skype will simultaneously translate as you're talking to somebody. These are all major AI achievements over the last few years. Of course, we had IBM's Watson defeat the world champion in "Jeopardy." I should point out that that's IBM Watson, the program, not IBM Watson, the brand that came later. That's a whole other story. There really is a lot of hype there. Of course, there is success in robots. My point is there are a lot of impressive things going on. What's really interesting is that all these successes come from a very simple paradigm that's been remarkably successful.

This is probably my most technical slide. I want to show you how all this AI success is generated, even if you're not a computer scientist. It really comes from something called machine learning. I should caution you — our field, artificial intelligence, machine learning, is full of grandiose labels. Artificial intelligence is still not quite intelligence. Machine learning isn't exactly learning.

Let's talk about what it does do. Machine learning is, you take a bunch of categories pre-specified by a person. In my example, we have some cats. We've got some dogs. We've got some flying animals. Those are the categories. You take some examples of these categories. These are typically represented as vectors of numbers, and you associate a label with each vector. The vector says this is a cat. This is a dog.

When I say *you*, I should be clear. This is done manually. The computer is not doing that. The definition of the categories, the creation of the labeling of the data, is all done by people. There's a machine learning algorithm that does what I like to call the last mile of learning. After everything has been set up with copious manual labor by people, then a statistical algorithm looks at the labeled data and creates a statistical model that is then used to make predictions.

The next vector I see is a cat. The next image I see is a dog. This is the paradigm of machine learning. What's remarkable is that you can apply this paradigm to speech signal, you can apply it to email messages, categorize them as spam or not. You can apply it to road following, cars staying on the road — this is what's been driving all the AI successes that you've heard of. This is great stuff. If you ask me what's going to happen over the next five or ten years, we're going to see more and more of this in lots of arenas, in healthcare, in enterprise, in all kinds of places.

At the same time, we have to remember that this is still a very limited technology. These capabilities are very, very narrow. If we can do chess, we can't do other things. The program that plays Go can't do other things. We'll talk about that. I like to call these things AI savants, because they're so, so narrow in their capabilities. The poster child of this — on the one hand, tremendous but at the same time narrow capabilities — is, of course, AlphaGo, the program that in March 2016 defeated the world champion in the ancient game of Go, which is a fantastically difficult game with many, many options.

The thing is it's still a very, very limited system. I imagine to myself, let's say, I go to some AI cocktail party, the kind of place where geeks go. Let's say I meet AlphaGo and we were to have a little dialogue. I might say 'AlphaGo, congratulations! You defeated Lee Sedol, but can you play poker?' AlphaGo would say, 'No. I can't.' I'd say, 'Can you cross the street?' AlphaGo would say, 'No. I can't do that either.' I'd say, 'Can you at least tell me something about the game of Go,

the game that you won?' AlphaGo would say, 'No.' It can't explain itself. The remarkable thing is AlphaGo doesn't even know it won. All it is is a fancy calculator applied not to multiplication and division but to analyzing the different possibilities in the game of Go, which, of course, is a game that's black and white with moves. It's very, very artificial and stylized.

What we need to understand is programs like AlphaGo are intelligent in a sense, a narrow sense — it's a savant — but what they *aren't* is autonomous. This distinction is really, really important because we're used to thinking of intelligence and autonomy as going hand in hand. The intelligence that we're familiar with is intelligence of people, and people are, of course, autonomous as well. With machines, it's really, really different. To make this distinction clear, I'm setting up a two-by-two table so we can analyze this. What I'm trying to show you here is that intelligence is not autonomy. These are orthogonal notions.

"It's very scary to have a weapon that can make a life or death decision without a human in the loop. The intelligence is not the problem. It's the autonomy."

Let's look at the upper left-hand corner here. What we have is high autonomy with low intelligence — that's a bunch of teenagers drinking on a Saturday night. In contrast, on the bottom right, we can have high intelligence with very low, minimal, essentially no autonomy. That's a program like AlphaGo. Its autonomy is restricted to choosing which move to make in the game. It can't even decide on its own to play another game. That's how limited it is. This distinction is really important.

Let's look at various other examples and remember it. Let's take, for example, computer viruses. They're dangerous, potentially very destructive, but their danger comes from their autonomy, from the fact that they can go from one machine to the other over the internet, duplicate themselves, potentially causing massive damage. They're not intelligent.

I like to point out that my seven-year-old is really more autonomous than any AI system. He's very intelligent too, but that's not my point. He can cross the street. He can speak English. He can hear English, at least when he wants to. He doesn't play Go, but he's a highly autonomous being — again, the opposite of this savant-type behavior we see in AI systems.

Let's talk about AI weapons. That's a topic that has received some attention. Often it gets people's heart racing. You start thinking about 'Do I really want intelligent weapons?' Imagine a weapon that can launch itself, fly halfway across the world, and kill somebody. That's the stuff of nightmares.

What I want to point out is that the nightmare has to do with the autonomy. It's very scary to have a weapon that can make a life or death decision without a human in the loop. The intelligence is not the problem. It's the autonomy. Intelligence in weapons could actually prevent mistakes like we've had where innocent civilians get killed. So, the key thing that we want to avoid, again, is not intelligent weapons but autonomous ones, or at least we want to think about very carefully. Again, I'm just giving these examples to set up this key distinction.

To those of you who are very optimistic about AI, I do want to pose the question about what's AI's impact going to be on jobs? That's a very serious question that merits a lot of discussions that I don't necessarily have the answer for. Hal Varian, the chief economist at Google, has said that old jobs are going away but new jobs are going to come and replace them. I, frankly, don't think it's that simple. It's a question of, at what rate? Old jobs have gone away. New jobs have come. But it seems like this change is happening at an unprecedented rate.

I'm not a Luddite. I'm not suggesting that we just stop AI. These are pictures of the Luddites throwing shoes into the looms because they were concerned about their jobs in the textile industry. Obviously, the progression of technology has a lot of benefits. Think about washing machines, antibiotics, textiles — it's not something that we can just rule one way or another. Still, people ask the question, 'If it's not clear what impact AI will have on our society, if there are pluses and minuses, why don't we at least declare a moratorium on AI? Why don't we slow down and give ourselves a chance to think about it?' That can be quite appealing.

"... if we slow down AI progress in this country, we very much do so at our peril. Right now, we have something of an edge."

Bill Gates himself said — when you come from Seattle, you say things like 'Bill Gates himself.' This is almost like a deity. Bill Gates is a very smart guy. He said, 'Why don't we put a tax on robots?' Which, of course, has the effect of slowing things down, at least the proliferation of robots. The problem I have with that idea is that AI is very much a global phenomenon. China has declared explicitly that they want to be the world leader in AI by 2030. It's right around the corner. Putin has said that the leader in artificial intelligence will rule the world.

So, if we slow down AI progress in this country, we very much do so at our peril. Right now, we have something of an edge. I'm not sure that we want to give that up. Even in the case of weapons, which is a very, again, tricky issue, I don't mean to simplify it. But when people ask me about AI weapons, I say the one thing that I fear more than highly powerful AI weapons in the hands of our

military is highly powerful AI weapons in the hands of rogue nations, in the hands of terrorists. We actually benefit from a healthy competition in this area.

There's a whole other dimension to this, though, and that's the dimension that Paul Allen, my boss, had in mind when he created the Allen Institute for Artificial Intelligence. That's to set up a mission of AI for the common good. So, at Allen AI, or AI2, as we call ourselves, we're not trying to use AI to create weapons, to violate people's privacy, not even to sell you things. We're really trying to use AI to make the world a better place. I want to give you two key examples just to highlight the potential benefits of AI; one example we're working on, and the second example we aren't, but many other people are.

The first example has to do with AI-based scientific breakthroughs. The number of papers — and I think I don't need to tell anybody in this audience this fact — the number of scientific papers is growing explosively. It's actually doubling every few years or so. There are thousands and thousands of papers on cancer alone being published every week. Let me tell you, they're not getting easier to read, either. The number of papers that any of us can read in our lifetime is relatively small, and the number of papers is growing. So, we have a problem. There are no Renaissance men and women anymore. What we need is some kind of tool to help us be better scientists, to help us be better engineers.

We have a project at the Allen Institute of AI called Semantic Scholar. What it's doing is using natural language processing, the ability for computers to understand certain things in text — not understand Shakespeare and nuance connotations but try to extract the basic facts from turgid scientific prose. Map that into knowledge that the computer can use and that scientists can use. They don't have to read all these thousands and thousands of papers.

"'What if the cure for an intractable cancer is right now hidden in all these thousands, if not hundreds of thousands, of different papers and different studies? Can machines help us by teasing apart this information, synthesizing it, and presenting it to a medical researcher to help him or her make progress on that cancer?'"

Without going into the technical details here, our vision is to say, 'What if the cure for an intractable cancer is right now hidden in all these thousands, if not hundreds of thousands, of different papers and different studies? Can machines help us by teasing apart this information,

synthesizing it, and presenting it to a medical researcher to help him or her make progress on that cancer?' That's some of the potential that we're working on every day at the Allen Institute for AI.

My colleague, Eric Horvitz, likes to say it's the absence of AI technologies that's already killing people. It's not that AI is going to be terminal-like and kill us. It's quite the opposite. He's not just talking about Semantic Scholar and the information in scientific text. The third-leading cause of death in American hospitals is some kind of doctor error. Information systems, AI systems that can analyze what's happening in a hospital, that can detect potential mistakes that exhausted and overworked doctors make, could really save an enormous number of lives.

A whole other arena, this is one that we're not working on, is driving. Frankly, human drivers worry me. I have a 17-year-old. He's texting and driving. I know this because he's in the car and I get a text from him. I have this moral dilemma — do I text him back and say stop texting or do I not text because I don't want him to read my text? You worry about your kids. They're texting and driving. Even if your kids aren't texting and driving, other people's kids are and they're going to run into your kids. It's a huge problem. Certainly, some of his classmates are drinking and driving. DUI is a major problem for us.

It doesn't even have to be this nefarious. A while ago a friend of mine in Seattle was jogging and he was hit by a car. Thankfully, he's okay, the car was going very slowly. It was driven by a 96-year-old driver. It was not a great idea, but, at the same time, as we get older and we have parents, we don't want to limit their mobility. We want them to continue to be independent and so on, but how do you do that and still keep the rest of us safe? Obviously, self-driving technology can make a huge difference here, a very positive one. There are more than 30,000 highway deaths each year, close to a million accidents. A lot of these can be prevented if we have better technology using AI techniques.

Again, remember, AI is a tool here. It's really not that different than anti-lock brakes or automatic transmission. It's not like these cars that are going to become increasingly safe. It's not like a hundred cars are going to band together and say we're going to take over the White House. They don't decide where to go. They're tools at our disposal. It's important to remember that, again, this distinction between intelligence and capability, safe driving and autonomy, which still remains in our hands.

The question then becomes, 'Am I suggesting that everything's fine? AI's going to be beneficial. Sure we have to worry about jobs, but we'll leave that to the economies. But let AI blossom unconstrained, unfettered.' No, I'm not saying that either. It's a complex issue. We need to think

about how we prevent AI from harming us. What are negative uses of AI? When you think about that, there's actually been relatively little written or thought about this. One almost naturally goes to Asimov's Three Laws of Robotics. How many of you have read some of Asimov's stories? Okay, most of you but not everybody.

Let me quickly review, because even for you it's been a long time. His three laws are: 1) A robot may not allow a human being to come to harm — it seems like a good idea; 2) A robot must obey its orders so long as it doesn't conflict with the first law; so you can't order a robot to harm somebody; and then 3) A robot should protect itself, again, so long as it doesn't conflict with the first or second laws. These laws are really quite elegant, and they've survived for more than 60-70 years. At the same time, they're quite ambiguous.

"... to understand what's harmful or what's not really requires common sense. Remarkably, that's been one of the hardest things for us to give to the machine."

Asimov had all these stories that you read where he showed that there's contradictions. There are problems. It's not simple to enforce or even understand what these laws are. It has a lot to do with, what is this notion of harm? You can say "harm," but how do you communicate that to a computer? All the computer understands is a programming language.

Imagine that I have Alexa on my laptop, or Cortana, or whatever it is. I tell it, 'Reduce utilization on my hard drive.' It says, 'Yes, Oren, done.' I say, 'What did you do?' It says, 'I took your dataset that took you years to assemble. It's not backed up anywhere. There's about 15 gig. I deleted it. That proposal you wrote for funding. I deleted that too and all the copies. I've saved you a lot of space.' So, it was obeying my command but in doing that, it created — I see people shuddering. It creates a huge amount of harm.

The problem is that to understand what's harmful or what's not really requires common sense. Remarkably, that's been one of the hardest things for us to give to the machine. There really are no machines today with even a modicum of common sense. I like to represent that with what I call the AI car wash. This is a guy washing his car in the rain. It's not a great move. This is what AI systems are like today. While they can play Go very well or even do speech recognition, they have no common sense. That's a huge problem if you want them to be able to do things and, at the same time, avoid harm.

Basically, Asimov's laws are fantastic, but they're not very practical. In trying to be more pragmatic about it, I wrote an Op-Ed piece for *The New York Times*. It just came out in September. I tried to

suggest a regulatory framework. Again, not all the answers by any means but some ideas for thought. How can we think about constraining AI? First of all, to those of you who raised your hands being scared of AI, I do believe that we ought to put an impregnable "Off" switch on any AI system. This is a picture from *2001: A Space Odyssey* where HAL, the computer, says, 'I can't do that, Dave.' When the computer is killing Dave and says, 'I can't do that,' we need to be able to just turn the computer off. That's a fundamental principle.

"The line between computer technology and AI technology is actually very unclear. Instead, my suggestion is, let's regulate AI applications. Let's not try to regulate the research."

Another thing, too, to think about is the fact that it is very hard to regulate or constrain the research field itself. It's fast moving. It's amorphous. The line between computer technology and AI technology is actually very unclear. Instead, my suggestion is, let's regulate AI applications. Let's not try to regulate the research. When we have cars that have AI and, of course, have the potential to cause accidents, when we have toys that have AI and have the potential to violate privacy, when we have robots in the workplace and potentially in the home, those are things that we can regulate, and we should.

"We should have a rule that a computer system should engage in full disclosure and disclose that it's an AI system ... we don't want our AI that's privy to more and more information about us to reveal that information."

What might that look like? The first thing we have to adopt is the notion of responsibility. An AI system ought to be subject to all the laws that apply to its human operator, its human manufacturer. If my AI car crashes into yours, I can't say, 'Don't blame me. It was my car.' 'My AI did it' is not an excuse. We have to take responsibility for our intelligent cars the same way we have to take responsibility for our unintelligent cars. That's an important legal principle. It needs to be elaborated, but it will help prevent irresponsible use of AI.

A second thing that's becoming increasingly important is full disclosure. It's easier and easier. That's going to change even more so in the future for an AI to pretend that it's actually a person. We see this in Facebook. We see the Twitter bots. We now found out that in the most recent election there was more and more of bot activity masquerading as human activity, etc. We should have a rule that a computer system should engage in full disclosure and disclose that it's an AI system.

Another remarkable thing that's happening is our technology is violating our privacy more and more. That's even before we have AI. This is from a recent article. It turns out that, in certain instances, Google actually gives a user a number of options, including saying, 'I don't like this ad because it knows too much about me.' This is not science fiction, this is not a proposal, this is a real thing that Google has unveiled because some people are just getting more and more upset by how much our technology knows about them.

If you think of systems like the Amazon Echo, which records audio in your house. AI Barbie — we have these Barbie dolls with chips in them engaging in dialogues with our kids. Who knows what information our kids are telling Barbie? I find that more scary than funny. Even the Roomba — you think, 'I've got this little hockey puck robot. It's cleaning dust in my house. What's the big deal?' It turns out that in the process of doing that, it's building a map of your house. Apparently, iRobot, the company that manufactures this, was actually considering selling this information to third parties. They didn't do it, but they were considering it. It's not something that you thought that your AI robot was doing.

"We want our AI to avoid bias. AI actually has a great potential to catch human bias, whether it's in judges or in people to alert us to, say, loan processing gone awry."

Imagine more sophisticated robots. They pick up the phone and say, 'He needs a new carpet. It's really fraying on the edges. Help me out here.' This is not something you want happening without your approval. So, we don't want our AI that's privy to more and more information about us to reveal that information.

Another topic — this I have to confess wasn't covered in my Op-Ed, partially for space reasons — I'm adding one more regulatory principle. We want our AI to avoid bias. AI actually has a great potential to catch human bias, whether it's in judges or in people to alert us to, say, loan processing gone awry. We have studies that showed that when judges get hungry, they tend to produce more negative decisions. So, AI could alert to, 'Your Honor, maybe it's time for a snack, your blood sugar is dropping.' There's a lot of benefit here.

There's a really interesting problem. This is a bit technical with AI. It goes like this. The data that we give our machine learning system is typically culled from the real world. It may contain some bias in it, all kinds of bias. What's remarkable about machine learning technology is that it tends to generalize. It attempts to compress the data into some general principle. When it does that, it can actually amplify the bias that's in the training data, which is very negative. So, whatever bias we have in the training data, let's say loans were denied to people of a certain gender or a certain

race some percentage of the time, the last thing we want is AI to automatically say, 'I get the pattern here. Let's be more aggressive on denying that. That's what my data is telling me.'

This, by the way, is a problem that does have a technical answer. One of the research scientists at AI2 just won a best paper award on work to avoid amplifying bias even beyond what's in the training data. That's another important issue. We would agree, we want to avoid bias.

I could go on and on, but I want to leave us time for questions and discussion. I just want to conclude by highlighting what I consider the most important point. I started with this question: Is AI good or evil? My answer is it's neither. It's neither good nor evil. It's a tool. It's a technology. More than anything, it's a pencil. It's a fancy pencil, one that we can draw amazing pictures with. But a pencil is a tool that we use. We get to choose. Do we draw nice pictures, or do we draw unpleasant, horrific pictures? The choice is ours. I hope you'll join me both in the conversation and in working as a society to make sure that AI is used for good, not for evil.

Thank you very much.

"Is AI good or evil? My answer is it's neither. It's neither good nor evil. It's a tool. It's a technology. More than anything, it's a pencil. It's a fancy pencil, one that we can draw amazing pictures with. But a pencil is a tool that we use."

Mr. HURD. Dr. Buck, you're on the clock, 5 minutes.

STATEMENT OF IAN BUCK

Mr. BUCK. Thank you, Chairman Hurd, Ranking Member Kelly, and distinguished members of the committee. I appreciate your invitation to give testimony today on this important subject of AI.

My name is Ian Buck. I'm the vice president and general manager of Accelerated Computing at NVIDIA. Our company is headquartered in Silicon Valley and has over 11,000 employees.

In 1999, NVIDIA invented a new type of processor called the graphics processing unit, or the GPU. It was designed to accelerate computer graphics for games by processing millions of calculations at the same time.

Today, GPUs are used for many applications, including virtual reality, self-driving cars, AI, and high-performance computing. In fact, America's fastest supercomputer, at Oak Ridge National Labs, uses 18,000 NVIDIA GPUs for scientific research.

Our involvement with AI began about 7 years ago, when researchers started using our processors to simulate human intelligence. Up until that time, computer programs required domain experts to manually describe objects or features.

Those systems took years to develop and many were never accurate enough for widespread adoption. Researchers discovered that they could teach computers to learn with data in a process we call training.

To put that in context, to teach a computer how to accurately recognize vehicles, for example, you need about 100 million data points and images and an enormous amount of computation. Without GPUs, training such a system would take months. Today's GPU-based systems can do this in about a day.

The world's leading technology companies have aggressively adopted AI. Google and Microsoft's algorithms now recognize images better than humans. Facebook translates over 2 billion language queries per day. Netflix uses AI to personalize your movie recommendations. And all those systems rely on thousands of GPUs.

My job is to help companies like these bring intelligent features to billions of people.

But AI's impact isn't just limited to tech companies. Self-driving cars, as was mentioned, surgical robots, smart cities that can detect harmful activities, even solving fusion power, AI holds the best promise to solve these previously unsolvable problems.

Here's a short list of problems for which I think AI could help.

First, cyber defense. We need to protect government data centers and our citizens from cyber attack. The scale of the problem is mind-boggling, and we're working with Booz Allen Hamilton to develop faster cybersecurity systems and train Federal employees in AI.

Second, as was mentioned, healthcare. Nearly 2 million Americans die each year from disease. We could diagnose them earlier and develop more personalized treatments. The National Cancer Institute and Department of Energy are using AI to accelerate cancer research.

Third, waste, fraud, and abuse. The GAO reported that agencies made $144 billion in improper payments in fiscal 2016. The commercial sector is already using AI to reduce such costs. PayPal uses AI to cut their fraud rate in half, saving billions. And Google used AI to lower the cost of its data centers by 40 percent.

Fourth, defense platform sustainment costs. Maintenance costs are a huge challenge for the DOD, typically equaling 50 percent or more of the cost of a major platform, totaling over $150 billion annually. GE is already using AI to detect anomalies and perform predictive maintenance on gas turbines, saving them $5 million per plant each year.

These are complex problems that require innovative solutions. AI can help us better achieve these results in less time and at lower cost.

For the role of government, I have three recommendations.

First, fund AI research. The reason we have neural networks today is because the government funded research for the first neural network in 1950. America leads the world in autonomous machine vehicle technology because DARPA funded self-driving car competitions over a decade ago.

While other governments have aggressively raised their research funding, the U.S. research has been relatively flat. We should boost research funding through agencies like the NSF, NIH, and DARPA. We also need faster supercomputers, which are essential for AI research.

Second, drive agency adoption of AI. Every major Federal agency, just like every major tech company, needs to invest in AI. Each agency should consult with experts in the field who understand AI and recruit or train data scientists.

Three, open access to data. Data is the fuel that drives the AI engine. Opening access to vast sources of data available to the Federal Government would help develop new AI capabilities so we can eliminate more mundane tasks and enable workers to focus on problem-solving.

In closing, AI is the biggest economic and technological revolution to take place in our lifetime. By some estimates, AI will add $8 trillion to the U.S. economy by 2035. The bottom line is we cannot afford to allow other countries overtake us.

And I thank you for your consideration. I look forward to answering your questions.

[Prepared statement of Mr. Buck follows:]

NVIDIA Testimony
February 14, 2018

Subcommittee on Information Technology of the Committee on Oversight and Government Reform

Thank you, Chairman Hurd, Ranking Member Kelly, and distinguished members of the Committee. I appreciate your invitation to give testimony today on the important subject of A.I.

My name is Ian Buck. I am vice president and general manager of accelerated computing at NVIDIA. Headquartered in Silicon Valley, NVIDIA is one of the world's leading computer technology companies, with more than 11,000 employees.

NVIDIA invented a new type of processor in 1999 called the graphics processing unit, or GPU. It was designed to accelerate computer graphics for games by processing millions of calculations at the same time. Today, GPUs are used for a variety of applications including virtual reality, self-driving cars, AI and high performance computing. America's fastest supercomputer, at Oak Ridge National Laboratories, uses 18,000 NVIDIA GPUs for scientific research.

Our involvement with AI began about 7 years ago when researchers started using our processors to simulate human intelligence.

Before AI, programmers had to write complex software by hand. Those traditional computer vision systems took years to develop and were not highly accurate.

In 2012 a brilliant graduate student named Alex Krizhevsky, at the University of Toronto, trained his computer to automatically recognize and classify objects. He did that by processing one million images on two NVIDIA processors. Training his A.I. model took about a week; without our processors, it would have taken nearly a year. His results were so accurate, he won a competition against researchers who had devoted their careers to hand-coding systems. That moment was the Big Bang of modern AI.

AI systems require enormous computation. To teach a computer how to accurately recognize vehicles, for example, you'll need about 100 million example images of cars, trucks, buses, emergency vehicles, etc. Without GPUs, training a system to recognize those images would take months. Today's state-of-the-art systems can reduce this to within a day.

The world's leading technology companies have aggressively adopted A.I. Google and Microsoft algorithms can now recognize images better than humans, and can automatically tag and search photos. Facebook translates over two billion language queries a day. Twitter understands the live video content in social media and can learn to detect harmful videos. Netflix uses AI to personalize your movie and TV show recommendations. All those systems rely on thousands of GPUs. My job is to help companies like Amazon, Google, Facebook, Microsoft, IBM and others bring intelligent features to billions of people.

But AI's impact isn't just limited to tech companies. Self-driving cars. Surgical robotics. Smart cities that can detect harmful activities. Solving fusion power. AI holds the best promise to solve these previously unsolvable problems.

To solve a big problem with A.I., you basically need 3 things: massive amounts of data, massive amounts of computing, and talented data scientists.

Here's a short list of problems for which AI could help:

1. **Cyber defense.** We need to protect government data centers, our institutions and our citizens from cyberattack. The scale of the problem is mind-boggling. The CEO of Juniper Networks said, "It's a challenge that is ultimately beyond human capability. No human can keep up with the pace of software changes today." We're working with Booz Allen Hamilton to develop faster cybersecurity systems and train federal employees in AI.

2. **Healthcare**. Nearly two million Americans die each year from disease. We should diagnose them earlier, and develop more personalized treatments. The National Cancer Institute and Department of Energy are using AI to accelerate cancer research.

3. **Transportation.** Congestion cost U.S. drivers over $300 billion last year, according to a recent transportation study. We're working with more than 300 companies to develop autonomous vehicles to make our roads safer and more efficient. The technology exists to make road, rail and air travel much more efficient.

4. **Waste, fraud, and abuse**. The GAO reported that agencies made $144 billion in improper payments in fiscal 2016. The commercial sector is using AI to reduce such costs. PayPal used AI to cut their fraud rate in half, saving billions. Google used A.I. to lower the cost of cooling its data centers by about 40%.

5. **Defense platform sustainment costs.** Maintenance costs are a huge challenge for the DoD, typically equaling 50% or more of the cost of a major platform. Sustainment costs

for the Army, Air Force, and Navy are about $150 billion, according to a report from the office of the Undersecretary of Defense. AI could help lower these costs by identifying maintenance issues earlier. G.E. is doing just that: using AI to detect anomalies and perform predictive maintenance on gas turbines, saving $5 million per plant each year. The Air Force is working on an experimental program to gather and analyze mission data so that technicians can address issues before they become serious.

These are complex problems that require innovative solutions. A.I. can help us achieve better results in less time and at lower cost.

Role of the Government
I have 3 recommendations for how the federal government can help.

1) Fund AI research
The reason we have neural networks today is because our government funded research for the first neural network, in the 1950s. America leads the world in autonomous vehicle technology because DARPA funded self-driving car competitions over a decade ago.
While other governments are aggressively raising their research funding, U.S. government research has been relatively flat. We should boost research funding through agencies like the NSF, NIH, and DARPA. We also need faster supercomputers, which are essential for AI research.

2) Drive agency adoption of A.I.
Every major federal agency -- just like every major technology company -- needs to invest in A.I. Each agency should consult with policy advisors who have domain expertise and understand the benefits of AI, and each agency needs to recruit data scientists and AI experts.

3) Open access to data
Data is the fuel that drives the AI engine. The federal government has access to vast sources of information. Opening access to that data will help us get insights that will transform the U.S. economy. It would help American workers in the public and private sector by eliminating mundane tasks and enabling them to focus on problem-solving and applying creative solutions.

The OPEN Government Data Act, which the House passed in November 2017, is a great start. Government data is available, but much of it needs to be curated before it can be useful. These activities need funding for researchers to carefully analyze and curate datasets.

Closing

A.I. is the biggest economic and technological revolution to take place in our lifetime. By some estimates, A.I. will add eight trillion dollars to the U.S. economy by 2035. The bottom line is, we can't afford to allow other countries to overtake us.

Thank you for your time and consideration. I look forward to answering any questions.

Mr. HURD. I thank all of you.

Now it's a pleasure to recognize the gentleman from Kentucky for 5 minutes for his first line of questions.

Mr. MASSIE. To the doctor from Intel, I don't want to try to pronounce your name. Help me out with that.

Mr. KHOSROWSHAHI. Khosrowshahi.

Mr. MASSIE. Khosrowshahi.

You said that AI was aspirational, but now it's a reality. Where did we cross the threshold? In the '90s, I worked at the AI lab at MIT. I worked on the hardware, because the software problem was too hard. And it seemed like you could solve certain engineering problems in the software, but it still feels that way to me.

What milestone did we cross, what threshold?

Mr. KHOSROWSHAHI. So I hear this a lot, that people studied neural networks in the '90s and they're kind of curious what has changed. And so let me just put it into a broader context. The history of AI goes back to the 1930s. The individuals who started the field, John von Neumann and Alan Turing, they were also the first people to build computers.

So the history of AI and computing has been tightly intertwined. So computing, as Dr. Isbell mentioned, is really critical. Compute power has dramatically increased since your time to today.

Another, the next change is data. And the algorithms potentially have not changed so much. They might look very familiar to you. But there has been actually a remarkable amount of innovation in the space of machine learning, which is a dominant form of AI, and in neural networks that Ian mentioned that is the state of the art today.

And invariably, these things change with time. The state of the art in AI changes with time. But the three things that are different today are computing power, data, and innovation in algorithms.

Mr. MASSIE. This next question I'd like to ask all four of you.

If there were going to be an XPRIZE for AI, what is the next big milestone? What's the sword in the stone that somebody should try to pull out and if they do they deserve a big reward?

Dr. Etzioni.

Mr. ETZIONI. I would observe that every time we build one of these systems, whether it's in medicine or self-driving cars or speech recognition, we're kind of starting from scratch. We have to train them with these millions or hundreds of millions of examples. We have to set the architecture by hand, et cetera, et cetera, et cetera.

If we could build, as Charles was alluding to, more general systems, which is something that we're very far from being able to do today, a system that can work across multiple tasks simultaneously without being retrained by hand every time, that would be a major breakthrough.

Mr. MASSIE. So, Dr. Buck, what would it be for you? Maybe driving from New York to L.A.?

Mr. BUCK. I think we've had our XPRIZE in self-driving cars with the work that DARPA did to kick off the industry innovation. There's a huge market for the first car company to really come up with a mass-produced self-driving vehicle.

I think AI at this point has the opportunity to revolutionize individual fields, and some could benefit from an XPRIZE, certainly healthcare. I think if we can identify an opportunity to do personalized medicine, to look at the genomics data that we've been able to get flooded with, with new instruments, and apply AI to understanding the NED treatments that are going to solve diseases, many of them just need to be detected earlier. If we could find them early, we could treat them. If we wait until the symptoms surface with today's technology, it's sadly too late.

And if I had to add one more, I think there are huge opportunities for AI to improve our infrastructure, transportation, and just apply it to real modern problems today.

Kansas City is doing a great project right now on detecting potholes with AI. They're actually gathering all the data from the weather data, the traffic information, and trying to predict when a pothole is going to form on a particular road. They are now up to 75 percent accurate within about 5 to 10 feet. So they can go out there ahead of time and treat that road and tar it up before they have to tear it up to fix a pothole.

There are so many different applications of AI, I think those XPRIZES would be fun to watch.

Mr. MASSIE. Dr. Isbell.

Mr. ISBELL. So I think there's sort of two answers to this.

One, all of us have said in one form or another that AI is interesting in the context of a specific domain, and so there's an XPRIZE for every domain.

But the more general question, I think, the answer is in the AI lab from the 1990s. I was also in the AI lab in the 1990s, and my adviser was Rod Brooks. As you might recall, at the time he was building a system called Cog, and the goal of Cog was to build——

Mr. MASSIE. I remember Cog.

Mr. ISBELL. Yes. I was probably sitting in the back when he announced it with you.

The interesting thing about Cog was the idea was that they were going to build a 3-year-old. And I think that the general problem of intelligence is a difficult one, and the real XPRIZE is being able to build someone we would recognize as sophisticated as a 3-, 4-, or 5-year-old.

Mr. MASSIE. Okay. Just a speed round here, if you'll indulge me. All four of you, I'll start here on the left.

Since you mentioned the 3-year-old goal that Professor Brooks had, how far away is AI from passing the Turing test, the classic Turing test, where if you were talking to this being, sentient being in the computer, you wouldn't be able to recognize it as not a human? How many years away are we?

You go first.

Mr. KHOSROWSHAHI. Twenty-plus.

Mr. MASSIE. Twenty-plus.

Dr. Isbell.

Mr. ISBELL. I assume the day after I die, because that's how these things usually work.

Mr. MASSIE. Or the day after your funding runs out.

Mr. ETZIONI. I should caution that the Turing test as it's set up is kind of a test of human gullibility. I'm afraid that we'll pass it

much sooner than is said. But if your question is about true human-level intelligence, I agree it's 20, 25 years and beyond, effectively beyond the foreseeable future.

Mr. MASSIE. It's definitely easier to fool somebody than it is to convince them they've been fooled, right?

Dr. Buck.

Mr. BUCK. I agree with my colleagues. It's equivalent to worrying about the overpopulation of Mars at this moment.

Mr. MASSIE. But it's the question. So what's your guess?

Mr. BUCK. Oh, decades.

Mr. MASSIE. Decades. Okay.

Thank you very much.

Mr. HURD. The gentlelady from Illinois is recognized.

Ms. KELLY. Thank you.

A few of you talked about the investment that needs to be made in this and made into some of the agencies. So what amount of money per year do you think the Federal Government should invest in some of the science agencies and foundations that you were referring to? Because it's easy to say we should invest, but what's your realistic——

Mr. ETZIONI. None of us are a policy or budgeting expert, as you can see from the few seconds of silence, but——

Ms. KELLY. We're silent, too, so don't worry.

Mr. ETZIONI. Let me suggest that much more than China. We have a substantially larger economy. We should be investing a lot more.

Ms. KELLY. Do you know what China is investing?

Mr. ETZIONI. I don't know the exact numbers, but it's certainly in the billions, according to their recently released blueprint.

Ms. KELLY. Anybody else?

Mr. KHOSROWSHAHI. So I don't know the numbers exactly, but funding for NSF I think is on the order of billions. And this money is highly leveraged. And funding graduate students studying at AI universities is a really good way to spend the money to accelerate innovation in AI.

And we do this at our company. We invest heavily in university programs, many grad students, many labs. And we've seen a lot of return in this specific area. So money well spent.

So $3 billion versus $6 billion, the extra $3 billion will be hugely effective in spurring innovation in AI.

Ms. KELLY. I was going to ask you, since your company is big in this area, how are you spurring on diversity, more women, more people of color?

Mr. KHOSROWSHAHI. It is actually a prime directive that comes from our CEO. So it's something that he is very focused on. We have diversity requirements in our hiring. Everyone knows these requirements in our hiring process. We focus on it.

And in our field in particular, we've seen firsthand—I have—that additional diversity benefits in many ways. So we discuss bias, transparency, having diversity in the scientific demographics within our company. We have different ideas presented. Sometimes these issues that you brought up are highly nuanced and they surprise me.

And so, again, that's a directive from our CEO.

Ms. KELLY. Thank you.

Dr. Isbell, you talked about increasing diversity, but starting in K through 12. What do you think schools need to do K through 12 to spur interest or what resources do they have to have?

Mr. ISBELL. So two short answers to that. I'll answer the first one first.

They have to connect what AI and what computing can do to the lives of the people who are in school. That's the single most important thing.

One thing that you just heard is that every dollar you spend on AI has a multiplying effect. And it's true, because it connects to all these domains, whether it's driving or whether it's history, whether it's medicine, whatever it is. And just connect that what you're doing will help you to do whatever problem you want to solve.

But the main limiting factor fundamentally is teachers. We simply do not have enough of them. You asked me how much money you should spend. Whatever number you come up with, it's 10 times whatever you will come up with is the right answer.

But even if you spent all of that money, we are not going to be able to have enough teachers who are going to be able to reach enough tenth-graders in the time that we're going to need in order to develop the next-generation workforce. It simply isn't possible.

What we're going to have to do is use technology to make that happen. We're going to have to make it so that Dr. Etzioni can reach 10,000 people instead of 40 people at a time and can work with people who are local to the students in order to help them to learn. That's the biggest, I think, resource for bringing people in who are young.

Ms. KELLY. Thank you.

Mr. ETZIONI. May I just add something real quick?

It's not just the number of teachers, but it's teacher training. My kids went to fancy private schools in Seattle that had classes called tech, and I was really disappointed to learn that they were teaching them features of PowerPoint because the teacher did not know how to program. So we need to have educational programs for the teachers so that they can teach our kids.

And believe me, 8-year-old, 10-year-old, what a great time to learn to write computer programs. And it will also help at least with gender diversity and other kinds of diversity, because at that point kids are less aware of these things and they'll figure out, hey, I can do this.

Ms. KELLY. Also, we talked about not getting the immigrant community. I serve on the board of trustees of my college, and that's something that we talked about. And they shared that the amount of foreign students has gone down drastically, because they don't feel as welcome in the country, and it's in engineering and the STEM fields that that has happened.

So I think my time is about up. Oh, I can keep going.

One thing I wanted to ask, what are the biases you have seen because of the lack of diversity?

Mr. BUCK. I think biases are a very important topic. Inherently, there's nothing biased about AI in itself as a technique. The bias comes from the data that is presented to it, and it is the job of a good data scientist to understand and grapple with that bias.

You're always going to have more data samples from one source than another source. It's inevitable. So you have to be aware of those things and seek them out. And a good data scientist never rests until they've looked at every angle to discover that bias.

It was talked about in our panel, in our testimonies. The think I'd add is that an important part of it, to detect bias, is where did it come from?

Traceability is a term that's used a lot in developing AI systems. As you're going through and learning better neural networks, inserting more data, you're recording the process and development.

So when you get out to a production system, you can then go back and find out why did it make that incorrect judgment and find out where was that bias inserted in the AI process and recreate it.

It's very important for self-driving cars, and I think it's going to be important for the rest of AI.

If you don't mind me going back to your previous question, I also think it's important that the committee recognize that AI is a remarkably open technology. Literally anyone can go buy, on a PC, download some open source software. They can rent an AI supercomputer in the cloud for as little as $3 and get started learning how to use AI. There's online courses from Coursera, Udacity. Industry, too. NVIDIA has an industry program called the Deep Learning Institute to help teach.

So those technologies are remarkably accessible and open, and I think that goes to your diversity, making it available. It inspires students, kids with ideas of how they can take data and apply these technologies. There's more and more courses coming online. And I think that will inspire the next wave of AI workers.

Mr. ISBELL. If I can just add to that.

I think the first round of bias comes from all of our beliefs, including myself. The sort of fundamental thing we want to believe is that the technology is itself unbiased and must be and that it is no more biased than a hammer or a screwdriver. But we'll point out that both hammers and screwdrivers are actually biased and they can only be used in certain ways and under certain circumstances.

The second set of bias comes from the data that you choose, which is exactly what Dr. Buck said.

I'll give you an example. When I was sitting in an AI lab apparently across the hall from you, a lot of the original work in vision was being done, particularly in face recognition.

A good friend of mine came up to me at one point and told me that I was breaking all of their face recognition software, because apparently all the pictures they were taking were of people with significantly less melanin than I have.

And so they had to come up with ways around the problem of me. And they did, and got their papers published, and then they made better algorithms that didn't depend upon the assumptions that they were making from the data that they had.

This is not a small thing. It can be quite subtle, and you can go years and years and decades without even understanding that you are injecting these kind of biases just in the questions that you're

asking, the data that you're given, and the problems that you're trying to solve.

And the only way around that is to, from the very beginning, train people to think through, in the way that Dr. Buck said, to think about their data, where it's coming from, and to surface the assumptions that they are making in the development of their algorithms and their problem choices.

Mr. ETZIONI. Bias is a very real issue, as you're saying, as we're all saying. But we have to be a little bit careful not to hold our database system to an overly high standard. So we have to ask, what are we comparing the behavior of the systems to? And currently, humans are making these decisions, and the humans are often racist, they're often sexist. They're biased in their own way.

We know, you talked about the case with a judicial decision. We have studies that show that when the justices are hungry, you really don't want them to rule at that point. You want them to go to lunch.

So my perspective is let's definitely root out the bias in our systems, but let's also think about these collaborative systems where humans are working together with the AI systems, and the AI system might suggest to the person, hey, maybe it's time for a snack, or you're overlooking this factor.

If we insist on building bias-free technology or figuring out how to build bias-free technology, we're going to fail. We need to build technology and systems that are better than what we have today.

Mr. HURD. Ranking Member, we need an XPRIZE for that, you know, to figure out when I'm hangry and make better decisions.

Ms. KELLY. My last question is, those of you representing companies, do you have internship programs? How do you reach out into the community?

Mr. BUCK. Certainly. I think the most exciting work is happening in our research institutions and even at the undergrad and earlier levels.

We're a huge proponent of interns. Myself, I was an intern at NVIDIA when I started at the company and worked my way up to be a general manager.

So I'm a huge proponent of interns. They bring fresh ideas, new ways of thinking, new ways of programming. They teach us a lot about what our technology can do.

Mr. KHOSROWSHAHI. If I'm allowed to comment on your last question.

So we talked about bias, but this line of thinking applies to everything. So transparency. I heard accountability. Humans are largely not transparent in their decisionmaking. This is something that's been studied exhaustively by people like Daniel Kahneman.

So I think it's very interesting to hear this firsthand, but we have to be concerned about humans as well as machines. And when they interoperate, that's even more challenging.

But, again, humans are biased, humans are transparent. And this is something to be cognizant of in your decisionmaking. I just wanted to stress that.

Ms. KELLY. Thank you.

Mr. HURD. One of the reasons we do these kinds of hearings is to get some of the feedback from the smart people that are doing this.

And, Dr. Buck, for example, we continue to do our FITARA Scorecards looking at how the Federal Government implements some of these rules. One of the questions we're going to start asking our Federal CIOs is, what are you doing to introduce artificial intelligence into your operations?

So, Federal CIOs, if you're watching, friends at FedScoop, make sure you let them know that's going to be coming on the round six, I think, of the FITARA Scorecard.

Where to start? So, yes, basic research. It is important. What kind of basic research? Do we need basic research into bias? Do we need basic research into some aspect of neural networks? Like, what kind of basic research should we be funding to start seeing that, to raise our game?

And all these questions are open to all of you all, so if you all want to answer, just give me a sign, and I'll start.

But, Dr. Buck, do you have some opinions?

Mr. BUCK. Certainly. As data science in general becomes more important to understanding the root cause of bias and how it is introduced and understood, I think it is a very important basic research understanding.

A lot of this work has been done. It can be dusted off and continued. I think it will be increasingly important as AI becomes more of the computational tool for changing all the things that we're doing.

Industry will tackle a lot of the neural network design. You have some of the smartest people in the world here in the U.S. building newer, smarter neural networks. They're largely focused on consumer use cases: speech recognition, translation, self-driving vehicles.

I feel like the science applications of AI, how AI can assist in climate and weather simulations, how AI can assist in healthcare and drug discovery, are still early. And it is an area that has less of a commercial application but obviously really important to this country.

You have some amazing computational scientists at the DOE labs that are starting to look at this. I think they also recognize the opportunity that AI can assist in simulation or improve the accuracy or get to the next level of discovery. I think there are some real opportunities there.

And we're starting to see that conversation happen within the science community. Any more encouragement and, of course, funding to help amplify it would be greatly appreciated.

Mr. ETZIONI. I think you make a great point. There is the investment from Google, Intel, and Facebook. But there is so much basic research that they won't do.

And I also can't emphasize enough how primitive the state of AI is. Sure, we've made a lot of strides forward, but——

Mr. HURD. Not to interrupt, but give me some. What are examples of basic research they won't do that we should be doing?

Mr. ETZIONI. Common sense. Something that you and I and every person knows and AI does not. That a finger has five hands.

That people typically look to their left and their right before they cross the street.

There's an infinite set of information that machines don't have. As a result, they really struggle to understand natural language. So we've seen success where the signal is very limited, like in a game of Go or in speech recognition.

But all you have to do is turn to Alexa or Siri and realize just how little our AI programs understand and how little can we have a conversation with them.

So I think research into natural language processing, into commonsense knowledge, into more efficient systems that use less training data, all of these are very, very challenging fundamental problems. And I could go on and on.

Mr. HURD. Gentlemen.

Mr. ISBELL. So I have very strong opinions about this, but I will try to keep it short.

I think if I were going to pick one—I'm going to give you two answers—and if I was going to pick one thing to focus on that I don't think we're doing enough of, it is long-lived AI.

That is, a lot of the work that we're doing are systems that solve a specific problem for a specific relatively short period of time is why it ends up looking like supervised learning as opposed to something like long-term decisionmaking.

But if you think about what makes human beings so interesting, there are two things. One is that we depend upon each other, and the other is that we learn and we live for a really long time, not measured in minutes or hours but measured in decades.

The problem of reading is hard. It takes human beings 6, 7, 8 years to learn how to read. We need to understand what it means to build systems that are going to have to survive. Not just figure out how to turn the car now, but have to figure out how to live with other intelligent beings for 10, 20, or 30 years. That's, I think, a sort of truly difficult problem.

But having said that, I'll back off and say, I think the answer is you trust your agencies who talk to the community. NSF has a long list of things that they believe are important to invest in AI and other things as well and the get that by having ongoing communications and conversations with a large community. It creates a kind of market, as it were, of what the interesting ideas are.

And I trust them. I listen to them. I talk to them. They're the mechanism that sort of aggregates what people are believing.

And then, in some sense, what you can do or what government can do or what these agencies can do is to push us a little bit in one direction or another by giving incentives for thinking about a problem that people aren't necessarily thinking of.

But, in general, I trust the people who are doing the work.

Mr. HURD. Dr. Khosrowshahi.

Mr. KHOSROWSHAHI. So we've been talking about high-level aspects of AI, decisionmaking and so forth. But in some of our testimonies we mentioned that there is a substrate for computation that enables AI. You have lots of data, need a while to compute.

We're at an interesting point in time where we're having rapid innovation in AI, lots of successes. It's being driven by availability of data and compute. The amount of data is increasing really, real-

ly rapidly, and the compute has to commensurately increase in power.

So that will require basic research and innovation at the silicon level, at the hardware level, which is what Intel does. We have fabs. We build the hardware from glass.

So areas such as silicon photonics, analog computing, quantum computing, low-powered computing, all of these areas are potentially great investment NSF funding opportunities for you.

And I'd like to also mention the landscape for getting AI systems to work involves so many different things. It requires machine learning, teachers, and so forth. But it requires things that seem prosaic but are really important, reliable software systems that are accountable, scalable, robust, and so forth.

Again, that comes from investing in STEM and computer science in early stages of someone's career development.

Mr. HURD. So we've talked about bias as a potential challenge that we have to deal with as we explore and evolve in the world with AI. Another way you can manipulate a learning algorithm is by loading it up with bad data.

What are some of the other challenges and other threats to artificial intelligence that we should be thinking about at the same time that we think about bias and integrity of the data that's involved in learning? Anyone.

Dr. Buck.

Mr. BUCK. I'll emphasize that it's easy to say we have lots of data. It's actually quite challenging to organize that data in a meaningful way. The Federal Government has vast sources of data. It is very unstructured.

Mr. HURD. Very aware.

Mr. BUCK. And that is a challenge. We just spent a decade talking about big data. And as far as I can tell, we've largely collected data, not really done much with it.

You now have a tool that can take all that data you've collected and really have some meaningful insights, to make a new discovery in healthcare, to save enormous amounts of money by finding inefficiencies or, worst, waste or fraud. But that data needs to be aggregated, cleaned up, labeled properly, and identified.

I certainly would make sure that not only that the Federal Government has an AI policy but also has a sister data policy as well to organize and make that data actionable and consumable by AIs, whether within the Federal Government or make them available to the larger research community.

I am sure there are dozens, if not thousands, of Ph.D.'s waiting to happen if they just had some of the more interesting Federal data to really make those kinds of discoveries.

Mr. HURD. Well, Dr. Buck, one of the first things this committee looked at was the DATA Act. And, shocker, the Federal Government was actually ahead of the game in trying to make sure that we're taking on that data and adding some structure to it. Implementation of that, as you have pointed out, is a bit tricky. So any tools that you all have to help with that would be great.

Other concerns?

Dr. Isbell.

Mr. ISBELL. So I'll add one. I agree with everything that Dr. Buck said and what other people have said before. Data is the problem. But one real issue is we typically build AI systems that don't worry about adversaries.

So this ties back into the notion of long-lived AI systems. So we're building a system that's going to determine whether you have a tumor, whether you have a heart attack, whether you should get a mortgage, but we're not spending a lot of energy—some people are thinking about this—we're not spending a lot of energy figuring out what happens when we send these things into the wild, we deploy them, and other people know that they're out there and they're changing their behavior in order to fool them.

And how do we make them change over time is an arm's race. You can think about this security. It's easy to think of. We could think of something even simpler, like spam. I get all this terrible mail. I build a system that learns what my spam is. The people who are sending spam figure out what the rules are and what's going on there, and then they change what they do. And it just keeps escalating.

And so this notion that you're going to have to not just solve the problem in front of you but solve the problem as it's going to change on the next round, the round after that, and the round after that, I think that's a real limitation of the kind of way that we build systems, freeze them, and then deploy them.

And I'm not saying that that's all people do and that no one is thinking about it. But I do think, because we tend to think in this sort of a transactional way about AI, we sometimes don't think through the consequences of having long-term systems.

Mr. KHOSROWSHAHI. I'd like take a slightly different tone. So we have talked in our testimonies about bias, privacy, transparency, assurances of correctness, adversarial agents trying to take advantage of weaknesses in the system.

So one thing that I've seen in this past year that I haven't seen in the past 10 years is these things are discussed at academic conferences. Companies like Intel, my team, actually these are some of the top priorities, these issues that you raise. They're discussed. They're attracting some of the best minds in the field.

I just introduced the idea of transparency literally months ago. And it's a really interesting area. It's highly nuanced. Humans are a tribal, multi-agent society. There are times when, if people have more information, the overall performance of the system goes down. It's very nonintuitive. Things can happen. Academics are pouring a lot of effort into this area.

So I'm just very, very optimistic that the things we've enumerated today are being addressed, and we should just amplify them. So the government can play a big role in investing in things like academic research.

It is quite different to me—I don't know if you guys concur—but the last major machine learning conference, NIPS, was really eye-opening to me, that there is a workshop on transparency, there is a workshop on bias, there is a workshop on diversity in the demographics of the AI community.

So we are definitely on a very positive and virtuous track, and I'm asking government to just amplify this however it can.

Mr. HURD. The distinguished gentleman from the Commonwealth of Virginia is now recognized.

Mr. CONNOLLY. Thank you, Mr. Chairman.

And thank you to our panel.

Dr. Etzioni, from here, I had a little trouble reading what was underneath your name. And I thought for a minute it said alien AI. I thought, wow, we really are getting diverse in the panels we are putting together here. Alien AI.

Mr. ETZIONI. I come in peace.

Mr. CONNOLLY. Yeah. Thank God.

So we were reminded rather dramatically last September with the Equifax hack that compromised information on 145 million Americans as to the risks of devastating cyber attacks and the absolute need for creating shields and protective measures, both for the government and for the private sector.

According to the 2016 report from the NSTC, the National Science and Technology Council, AI has important applications in cybersecurity and is expected to play an increasing role for both defensive and offensive cyber measures.

Dr. Khosrowshahi—and I'm from now on going to say the doctor from Intel—how can AI be most useful in defending against cyber attacks?

Mr. KHOSROWSHAHI. So I'll suggest a few ways, and I guess we'll have other opinions.

So cybersecurity, of course, is a major issue broadly in computing, as well as in AI, and as well at Intel. It is one of our primary focuses.

So in terms of addressing cyber attacks using AI, cyber attacks are intentionally devious and nefarious, obscure. And these kinds of actions are really well suited to the latest state of the art in AI, machine learning.

That is algorithms can take large corpora of data—these are inputs from whatever the type of cyber attack you're experiencing—and they can build a model of the cyber attack and a response, essentially.

And the response can have very low latency. It can study the statistics of the attack, potentially it's a novel attack, build a model, and respond very quickly.

So that's one way we can address cybersecurity, is with better models to defend against it.

Another way—another thing that we can—it's not in answer to your question—but when we build models, it's good to know the set of possible attacks, because a researcher, a data scientist, is very cognizant of building robust models that are resistant to adversarial events.

So as we get knowledge of cybersecurity issues in this area, AI, we build in security and defense against cyber attacks into the models such that adversarial actions do not perturb or give erroneous results.

Mr. CONNOLLY. Presumably also one of the advantages of AI would be early detection. I mean, part of the problem of cyber, certainly from the Federal Government's point of view, but apparently in the private sector as well, is when we finally realize we have been compromised, it's too late.

Mr. KHOSROWSHAHI. That's right.

Mr. CONNOLLY. And AI has the potential for early detection and diversion, preemption, protective walls, whatever.

Mr. KHOSROWSHAHI. That's right. The nature of these attacks could be so devious that the smartest human security experts could not identify them. So can either augment our human security experts or we can have systems that are early detectors that can just flag this is a potential threat. And these systems are really well suited for doing this, latency and learning very quickly.

Mr. CONNOLLY. Anyone else on the panel is more than welcome to comment.

Dr. Etzioni.

Mr. ETZIONI. I just wanted to add that at the root of the Equifax hack was human error, several human errors. So something you might want to think about is, what are the incentives that we have in place to avoid that? What are the consequences that people at Equifax face—and not to pick on them—for making those mistakes with our data?

I think if we put the right incentive structure in place, it's not a technical solution, but it'll help people to be more watchful, and they should be.

Mr. CONNOLLY. Yeah.

Mr. BUCK. The statistics here are alarming. And the rate of attacks are growing exponentially way faster than we can expect a human operator, even with the tools they have today, to keep up.

This is a very hot topic in the startup community. There are many startups trying to apply AI to this problem. It's a natural fit.

AI is, by nature, pattern matching. It can identify patterns and call out when things don't match that pattern. Malware is exactly that way. Suspicious network traffic is that way.

One startup we work with, they're claiming the top AI software is only able to capture about 62 percent of the potential threats that are out there. But by applying AI, they can shorten the time to discovery and get to 90-plus percent accurate malware detection, and the false error rate, get it down to less than 0.1 percent where normally it's 2 percent.

It's an opportunity to increase the throughput of our detector systems and make them much more rapidly responsive.

Mr. CONNOLLY. So why aren't we doing it? Is it the cost?

Mr. BUCK. The AI just needs to be developed. It is in the process of being developed by those startup companies. It's not as talked about in application as maybe video analytics or ad placement, but it is certainly active.

Mr. CONNOLLY. Well, you put your finger on two things, among others. But one is the exponential growth in the volume of attacks. I talk to some Federal agencies, and I'm stunned at the numbers. I mean, I know of one Federal agency, not a big one, where the cyber attacks or attempted attacks are in the hundreds of millions a year.

And you're absolutely right. I mean, this particular agency, its mission isn't cyber. It's got a very human mission. And it's trying to put together through Band-Aids and other measures some protection. And it does raise questions about the ability of, in this case, the Federal Government to protect itself.

Mr. BUCK. I'm seeing a sea change in that as well. Not just are we looking to protect our firewalls and the data coming into our firewalls, but the data traffic behind the firewall.

Assume you are attacked, for the sake of argument, and look at the traffic that's inside your firewall to detect it. Because as was mentioned before, in many cases you may already be compromised and you don't know it.

So it's important to look at both, the front line as well as behind the lines, in understanding your network traffic and your security.

Mr. CONNOLLY. And the second thing this conversation I think underscores, and we had testimony yesterday from the intelligence community, but the idea that the Russians are not going to continue their attacks and attempts to distort our electoral process is naive. All 17 intelligence agencies in the United States Government testified to the fact that it is an ongoing threat and the midterm elections will be a target.

So in a democracy, that's the very heart of how we function. How do we protect ourselves? And I think maybe we've got one tool, maybe a very critical tool, in terms of artificial intelligence. But trying to get that out to the myriad localities, over 10,000 localities in the United States, is going to be a different kind of challenge.

I thank you, Mr. Chairman.

Mr. HURD. Mr. Lynch, you are now recognized.

Mr. LYNCH. Thank you, Mr. Chairman. I appreciate that.

Dr. Etzioni, in your written testimony you state, and I quote here, "We can and should regulate AI applications." Obviously, as more and more AI systems are used to collect more and more sensitive and personal data on Americans, there are palpable and real privacy concerns.

What are the ways in which you think that the regulations that you anticipate would serve to protect the private information of Americans?

Mr. ETZIONI. So I think that there are some principles that I can talk about. And, frankly, you and your staff are probably better qualified to think through specific regulations.

But a principle that I would really advocate is identifying when AI is involved. And that's something that we can regulate so that the bots, at least the homegrown ones, state that they're AI. We had Intel inside. We should have AI inside.

Most recently we've seen that there are examples of fake pornography, superimposed celebrities on top of bodies and things like that. If we can't trust the integrity of our pornography—obviously I'm joking.

Mr. LYNCH. Thanks for making that clear.

Mr. ETZIONI. But the point is we should label when AI is being used. And, likewise, we should be clear when we have AI systems in our homes. Alexa, AI Barbie, the Roomba vacuuming our floor, they naturally also vacuum up a huge amount of data, some of it from our kids, if Barbie is talking to our kids. We should have regulations about where that information can go.

Mr. LYNCH. So the proliferation of AI, I just see it proceeds at a velocity far exceeding the ability of Congress to keep up with it, and that's true with many technologies. And oftentimes we rely heavily on the private sector to look at those ways that, if AI is

being broadly used, how we might develop a protocol that would prevent that private information from just getting out there.

And we have, in a very narrow sense, the Equifax situation where we have the names, addresses, Social Security numbers of 150 million Americans out there, just gone. So they basically burnt the entire Social Security number system as a reliable and secure indicia. So that's gone. And it's just because one company was very lazy about protecting data.

And so I'm just concerned. I have similar concerns about AI being out there and these bots. And we've got some pretty creative hackers out there, Russians and others, that have been able to access some very, very sensitive information. At one point they swept every bit of data from any individual who had applied for a high-level security clearance in this country.

And so I could just see if there are, as you say, not necessarily household appliances, but other forms of AI operating a higher level, if those are hacked, it just increases the magnitude of our vulnerability exponentially.

And I'm just trying to think in advance, as this is all happening in real-time, how do we protect the people who elected us? We're all for innovation, but I think with the appropriate safeguards in place.

Mr. ETZIONI. The thing that I would like to highlight, though, is that you're right, those are some scary realities. But they are realties. They're often instigated from the outside. So maintaining our strategic edge.

And that's why I emphasize regulating applications as opposed to the AI field and AI research itself. If we adopt an overly defensive, dare I even say in a reactionary posture, we're just going to lose.

So this is a very competitive global business. And staying ahead, which we're all trying to do in various ways for education, et cetera, is essential.

Mr. LYNCH. Okay. Thank you.

I assume my time has expired, Mr. Chairman. I yield back.

Mr. HURD. Dr. Isbell, did you have a response to that question?

Mr. ISBELL. I just want to add something. I think it's important to recognize here everything that you brought up are deep concerns. But AI is a secondary problem there. The primary problem there is that we are sharing our data constantly.

Every one of you has a cell phone, possibly two of them, you have a watch, which is pinging all the WiFi hotspots everywhere you go. Each one of those devices has a unique ID. That unique ID is not you, but that unique ID is with you all the time. I can figure out with very little effort who you are, where you are, where you come from.

By the way, I've deployed systems myself, this is 10 or 15 years' old worth of technology, where I can predict what button you're going to press on your remote control after just observing you for one weekend.

We are creatures of habit. We are sharing our data in our cars, our phones, everything that we do. The data itself, even if it's anonymized, is giving amazing amounts of information about us as individuals. That's the primary problem.

The secondary problem is the AI, the machine learning, the technology, which can look at it very quickly and bring together the obvious connections even though you've tried to hide them.

But the first thing I think to think about is it's not the AI, because computers are just fast, that's just going to happen. It's the fact that we are sharing data, and we've given very little thought to what it means to protect ourselves from the data we are willingly giving to everyone around us. And I don't have an answer, but that, in some sense, is the root problem.

Mr. LYNCH. Mr. Chairman, if I could.

The ability of AI to aggregate the data, make sense of it, and give it direction and a purpose and a use, that's the magic of AI. The data's out there. And you're right, that's a problem. But I'm worried about weaponizing that raw data that's out there and how do we control that.

But thank you. I think you offered a very good clarification. Thanks.

Mr. KHOSROWSHAHI. Let me make a short comment.

So I liked to balance the discussion and present a slightly dissenting view to Dr. Etzioni.

Well-intentioned efforts, such as labeling robots and other devices that employ AI, it could have unintended consequences. You have in the State of California, my State, we now know that asparagus and coffee cause cancer. So we are going to have labels on every piece of food and every building that this thing causes cancer. And these signs are becoming uninformative.

So I would just be wary of unnecessary regulation or imposing regulation on a very young and rapidly moving field, because I can immediately see that it can have some adverse consequences.

We talked about transparency. To use your example, would you want something that is labeled and worse performing or unlabeled and better performing, to use your example.

And just in general, our view at Intel is that legislation should be based on principles, not on regulation that mandates certain kinds of technology. So we are self-regulating.

This field is wonderful, that it does a lot of high-minded academics who are now leaders in business, and there is a strong impetus to be good stewards of this technology to do good. And we have lots of things that we can impose on ourselves to self-regulate to potentially address some of the adverse conditions that you mentioned. Not all of them. Perhaps some of them do need legislation.

Mr. HURD. I've got some final questions. And this first question is for everyone. And I know you all have all spent your adult lives trying to answer this question, and so I recognize this before I ask.

And, Dr. Buck, I've got to give some kudos to your team that was out at the Consumer Electronics Show. They were very helpful in helping me understand some of the nuance of artificial intelligence. And if artificial intelligence was based on Fortran 77, I'd be your guy. That's my background experience.

But I understand how to introduce antivirus software into your system. I understand how you introduce CDM into a network. When we ask all the Federal CIOs how are you thinking about introducing artificial intelligence into the networks, the first question I'm probably going to get is, well, it's really hard.

And so my question is simple. And we've all been saying that AI is interesting because it's domain specific and I recognize how broad this question is. But how do we introduce AI into a network, into a system, into an agency?

Mr. BUCK. That's a great question. And AI can seem like rocket science. And first off, having this conversation is the first step. Explaining what it is and understanding it so they can comprehend it is, obviously, the first step.

And where I've seen it work most successfully is in meaningful simple pilot projects. Project Maiden, which is a project with DOD, where they're using AI to help with reconnaissance so that airmen are not staring at TV screens for 8 hours a day waiting for something to happen. They're letting the AI do the mundane parts of the job so our soldiers can do the decisionmaking.

That kind of application of AI is well established. People know how to do it. You don't need to invent a new neural network to do it. It's the same work that's being done elsewhere. But by creating these pilot projects inside of these agencies, they are dramatically improving the lives of the people that work there.

Mr. HURD. So do we believe we're at a point now where the agencies can be the ones that are involved in training the algorithm. Okay, you find an algorithm, you figure out what dataset you need to train it. And do you expect the person at Department of Interior to be the one training that, or is it folks that are providing that service?

Mr. BUCK. You can do it both ways. I've definitely seen public partnerships where agencies are going outside for consulting to help apply AI technology to a specific problem. Or in some cases the neural networks are well established. Image recognition is where AI started. It is a well-established technique. The networks are open source. The software is open source and public.

So I think if you find those use cases off the bat that are well published and, as was spoken, in these AI conferences well shared. The beauty about AI is that it's incredibly open, it's being done in the open source community, it's all being published. And it takes very little work to take one of those established workflows and apply it. And then the next step is to share that success.

Mr. HURD. Dr. Khosrowshahi.

Mr. KHOSROWSHAHI. So AI has changed over the last 80 years and it almost surely will change. We talked about neural networks. Five years from now, almost surely—I'm on TV—but I guarantee it's going to be something different.

But the underpinnings are you have data, you have model, you have inferences. You have data that has statistical distribution, whether it's images, whether it's a car driving down the road collecting video in the U.S. or Canada or wherever, different statistics. You build models, the models try to understand the statistics of the data, and then you can ask the model questions. Is this a cat or a dog? Is there a stop sign approaching me? That's basically what AI is today.

So if you just take these simple underpinnings and then apply them to whatever public policy or application CIOs want to insert into their business workloads and so forth, just understanding that basic element. There's going to be some data, it will have some sta-

tistical properties, maybe it will be difficult for a human to understand them. A machine could be better and faster, more robust, more power efficient than the brain. And then it can perform inferences.

And whether or not you choose to rely on these inferences depends on how good the model is, how much assurances of correctness you have. I mean, the landscape of AI is so vast and it's touching so many different things. And it's still, I would again stress, that it's very early on. We don't have artificial agents making decisions for us almost anywhere.

So even in finance, you would expect automated trading systems. It's not there yet. We're still in the very early stages. There is not widespread adoption in the industry. It will get there, but it's still early on.

But, again, the AI, the underpinnings and the applications, there's this model data inference. You can stick it in anywhere where that works.

Mr. ISBELL. So in the interest of time, I'll keep this short.

I want to distinguish between at least two different things. One is face recognition and that class of things versus shared decisionmaking. I think the answer for things like face recognition, relatively straightforward. At the risk of oversimplifying, it's like asking the question, how can we integrate the internet? How can we integrate telephones? It's relatively straightforward. It's well understood, it's very clear, and you can ask yourself how to use the screw driver.

The shared decisionmaking is what's difficult. That requires that the domain experts are part of the fundamental conversations. The research question from my point of view is figuring out how to be able to use humans in order to train the systems that we have when they don't understand machine learning and AI, but they do understand their domain. How do you get those people to talk to one another?

I'm not worried about the deployment of face recognition. I'm worried about how I'm going to get an intelligence analyst to understand enough about what it is they are doing so that they can communicate to a system that will work with them in order to make decisions.

That's where the difficult problem is, but it's really no different than just trying to understand what it is they actually do. The problem is, the thing that we know, is that people are terrible at telling you what it is that they do. You can't ask them and they tell you. You have to watch them, observe them, model them, and give them feedback. It's an iterative, ongoing process.

Mr. ETZIONI. I wonder if an approach would be to focus on outcomes and metrics and grand challenges. And if you ask for those rather than demanding AI and then they have to resort to AI to satisfy those mandates, that might work.

Mr. HURD. One minute for all four of you all to answer these two questions.

What datasets in the government do you want access to or should the AI community of people that are working on these challenges get access to? And what skill sets should our kids in college

be getting in order to make sure that they can handle the next phase when it comes to artificial intelligence?

Mr. ISBELL. All of them. And the skills that the students need in college, they need to understand computing. There shouldn't be a single person who graduates with a college degree who hasn't taken three or four classes in computing at the upper division level. They need to understand statistics. And they need to understand what it means to take unstructured data and turn it into structured data that they can construct problems around.

Mr. KHOSROWSHAHI. So on the datasets, things like NOAA, weather data, things that are not sensitive have private information, those would be the first. And there's a vast trove of this. This would be immediately useable by academics.

But on the skill set side, if I were to pick one, it would be computer science. I would invest as much as possible in teaching computer science K through 12, especially in high school.

Mr. HURD. Dr. Etzioni.

Mr. ETZIONI. Research funded by NIH, by NSF, DARPA, et cetera, is often not available under open access. Journals keep it behind pay walls. That's changing way too slowly.

So the dataset that I would like everybody, human and machine, to have access to is the data and the articles that you and we as taxpayers paid for. I think that's incredibly informant.

As far as the skill sets, I would say that everybody in college should be able to write a simple computer program and to do a simple analysis. And we can get there, and, remarkably, it's not required.

Mr. HURD. Dr. Buck, last word.

Mr. BUCK. I certainly would love to see all the datasets. I certainly also would like to see access to the problems around healthcare. And I know those are sensitive topics, but the problem is too important, the opportunity is too great, and it is where I feel like AI will truly save lives. If we could figure out to make that data available, it would be an amazing achievement.

In terms of education, I believe that data science is becoming a science again. And I also feel like training a neural network is not that hard. I think it can be done at the junior high level.

And the access to technology is available today. And I think we should start teaching students what this tool can do. Because it really is a tool and will inspire new applications that will come from the interns, the undergrads, the college students. That's what makes this fun.

Mr. HURD. Well, gentlemen, I think my colleagues would agree with me on this, this has been a helpful conversation. There is a lot packed into your all's testimony that's going to help us to continue to do our work on the Oversight Committee and to look at opening up some of these datasets. How do we double down on NSF funding? How do we focus on getting more? I think every kid in middle school should have access to a coding class. And we're working on that stuff down in the great State of Texas.

And many of these points that you make, we're going to be talking to folks in the government, in early March, in the second series of this AI series. We intended to invite GSA, NSF, DOD, DHS and to continue this conversation about how they are introducing and

looking at artificial intelligence and what more support they need from Congress.

So, again, I want to thank you all and the witnesses for appearing before us today.

The hearing record will remain open for 2 weeks for any member to submit a written opening statement or questions for the record.

And if there's no further business, without objection, the subcommittee stands adjourned.

[Whereupon, at 3:54 p.m., the subcommittee was adjourned.]